

**Encyclopedic
Dictionary
of
Mathematics**

Second Edition

**Encyclopedic
Dictionary
of
Mathematics**

Second Edition

by the
**Mathematical Society
of Japan**

edited by
Kiyosi Itô

**Volume I
A–N**

**The MIT Press
Cambridge, Massachusetts,
and London, England**

First MIT Press paperback edition, 1993

Originally published in Japanese in 1954 by Iwanami Shoten, Publishers, Tokyo, under the title *Iwanami Sūgaku Ziten*. Copyright © 1954, revised and augmented edition © 1960, second edition © 1968, third edition © 1985 by Nihon Sugakkai (Mathematical Society of Japan).

English translation of the third edition © 1987 by The Massachusetts Institute of Technology.

All rights reserved. No part of this book may be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from The MIT Press.

This book was set in Monolaser Times Roman by Asco Trade Typesetting Ltd., Hong Kong, and printed and bound by Arcata Graphics, Kingsport in the United States of America.

Library of Congress Cataloging-in-Publication Data

Iwanami sūgaku jiten (ziten). English.

Encyclopedic dictionary of mathematics.

Translation of: Iwanami sūgaku jiten.

Includes bibliographies and indexes.

1. Mathematics—Dictionaries. I. Itô, Kiyosi, 1915–. II. Nihon Sūgakkai. III. Title.

QA5.I8313 1986 510'.3'21 86-21092

ISBN 0-262-09026-0 (HB), 0-262-59020-4 (PB)

Contents

Forewords	vii
Prefaces	ix
Introduction	xix
Encyclopedic Dictionary of Mathematics	
Appendix A. Tables of Formulas	1421
Appendix B. Numerical Tables	1820
Statistical Tables for Reference	1837
Numerical Tables for Reference	1838
Journals	1840
Publishers	1848
Special Notation	1850
Systematic List of Articles	1857
Alphabetical List of Articles	1863
Contributors	1867
Translators	1873
Name Index	1875
Subject Index	1917

Foreword to the Second English Edition

The Mathematical Society of Japan takes pleasure in presenting this second edition of our *Encyclopedic Dictionary of Mathematics* to every researcher and user of mathematics. It is intended to be a compact, up-to-date source of information comprising, as completely as possible, all significant results in all fields of our Science, pure and applied, from the elementary to the advanced level. The success of the first edition owed much to the kind assistance given by the American Mathematical Society. As described in the preface, the members of our Society have taken responsibility for compiling this new edition. We hope that it will be as useful to the mathematicians of today as the first edition was to the mathematicians of yesterday. We also hope that this edition will be followed in years to come by subsequent ones incorporating the future development of our Science.

Hikosaburo Komatsu
President 1985–1987
Mathematical Society of Japan

Foreword to the First English Edition

The American Mathematical Society welcomes the publication of the *Encyclopedic Dictionary of Mathematics*. For many years we have been fascinated by the publication in Japanese, *Iwanami Sūgaku Ziten*, because we saw that this was an encyclopedia that contained effective and penetrating information about all the fields of advanced mathematical research. We were also frustrated because we could not read Japanese and so we could not really reach out to this expert and effective source of information. We now welcome the fact that the second Japanese edition has been translated into English and we look forward to the fascination which we can now have in getting at this rich mine of information.

Saunders MacLane
President 1973–1974
American Mathematical Society

Preface to the Second English Edition

This second edition of the *Encyclopedic Dictionary of Mathematics* is in substance an English version of the third edition of *Iwanami Sūgaku Ziten* (in Japanese). We shall explain how these two versions are related to each other and how the present edition differs from the first English edition. For the sake of simplicity we abbreviate *Encyclopedic Dictionary of Mathematics* and *Iwanami Sūgaku Ziten* to EDM and ISZ, respectively, and indicate the numbers of editions beyond the first by Arabic numerals.

The prefaces of the previous editions will clarify how ISZ, its augmented edition, ISZ2, ISZ3, and EDM came into existence in 1954, 1960, 1968, 1985, and 1977, respectively. EDM, published nine years later than ISZ2, consisted of its English translation and some new materials added to update its contents. In retrospect, it was felt that this procedure was not adequate to cope with the recent rapid progress of mathematics, and a suggestion was raised in the Mathematical Society of Japan that ISZ3 and EDM2 be produced simultaneously. The favorable reaction of the mathematical public to EDM encouraged us greatly.

In 1978 an agreement was made among the Society, Iwanami Shoten, Publishers, and The MIT Press for the publication of ISZ3 and EDM2. An editorial committee was established in the Society with the members named in the preface to ISZ3. Manuscripts were prepared simultaneously in Japanese and English by each contributor, with Yuji Ito acting as linguistic consultant.

I should mention that we benefited greatly from the kind comments on EDM by the following mathematicians: J. F. Adams, M. Atiyah, A. Borel, H. Cartan, K. Chandrasekharan, S. S. Chern, J. Dieudonné, E. Hewitt, F. Hirzebruch, O. Lehto, J. L. Lions, L. Markus, and J. P. Serre. In particular, we are deeply grateful to J. Dieudonné for his many detailed comments.

Compared with ISZ3 and EDM, EDM2 has the following characteristics:

(1) EDM2 contains many more new mathematical results than EDM. For the details we refer the reader to the main points of revision mentioned in the preface to ISZ3, because EDM2 has the same mathematical content as ISZ3 and the additions made to ISZ2 and subsequently in EDM are of relatively minor importance.

(2) The Japanese textbooks listed in the article references of ISZ3 are replaced in EDM2 by references to standard textbooks

written in English. Such replacement was not done in compiling EDM.

(3) Years of birth and death are included in the Name Index of EDM2 as far as possible. This information was given in ISZ2 but was removed in compiling EDM.

(4) The Subject Index of EDM2 is so designed that every concept consisting of two or more words can be traced from each of these component words. (This principle was adopted in ISZ2 and EDM but was not rigidly followed in ISZ3 due to the lack of space.)

(5) While all editions of ISZ were in one volume and EDM was in two volumes, EDM2 is in four volumes: I. Forewords, prefaces, introduction, text A–E; II. Text F–N; III. Text O–Z; IV. Appendices, indexes and other backmatter. The Systematic List of Articles appears in volume IV, page 1857.

While EDM2 is more voluminous than ISZ3, we hope that, being written in English and printed with generous margins, it can easily provide readers with information on every significant result of today's mathematics and so will be useful to a large number of mathematicians. I am responsible for any shortcomings that may exist despite all our efforts, and I would appreciate any remarks from the readers. I hope to have occasion to remedy any such shortcomings in the future.

Finally I would like to repeat my thanks to all the collaborators named in my preface to ISZ3. Also I express my gratitude to Yuji Ito for his tremendous effort to polish the English of all manuscripts, to Shigeru Iitaka for his laborious assistance in the final stage of the work, to Mrs. M. Nawata for her excellent secretarial work, and to the staff of the Department of Mathematics at Gakushuin University for providing me with an office for compilation for eight years.

Kiyosi Itô
December 1986

Preface to the Third Japanese Edition

The first edition of *Iwanami Sūgaku Ziten* was published in 1954, and a revised and augmented edition appeared in 1960. Extensive revision of the work was carried out subsequently, and the result was published in 1968 as the second edition, which has retained its valuable and useful life for the past 17 years. The English translation of this second edition published by The MIT Press has achieved international recognition. In the meantime, remarkable progress has been made in mathematical sciences: recognition of the interrelationships amongst various branches within mathematics has increased significantly, and the formulation of mathematics as a synthetic entity is in the making. Furthermore, advanced mathematical theories have been utilized frequently in the physical, biological, and social sciences, and expectations for mathematics to be the basis of all the sciences have been increasing. In order to cope with these developments, it was decided that the second edition of *Sūgaku Ziten* should be further revised and an updated version be published as the third edition. The English translation of this third edition will also be published by The MIT Press.

The main points of the revision are as follows:

(1) On the Size and Scope of the Encyclopedia: The total number of articles has increased only slightly from that of the second edition, from 436 to 450. However, in view of the rapid and extensive development in mathematics in recent years, many old articles have been reorganized and unified, and much new material has been added. Consequently, the mathematical content of the encyclopedia has increased a great deal over that of the second edition, and this, together with the authors' efforts to render the articles more readable than before, necessitated a 50% increase in the number of pages devoted to the main text. The account of computer science has been taken out of the subject area of Numerical Analysis, and a new subject area has been created for it; thus mathematics is now classified into 21 subject areas altogether instead of 20.

(2) On the Arrangement of the Articles: In the second edition, the title of each article was spelled out in Romanized Japanese, and articles were arranged alphabetically. In the third edition, however, the titles of the articles are arranged in the order of the Japanese syllabary (*katakana*) and each article is numbered

accordingly. To each article a parenthetical number is also attached, giving the subject area to which the article belongs and its place in that list.

(3) On the Text: Sections within each article are headed by letters A, B, C, ...; the initial section of each article gives an outline and is devoted to an introductory and general account of the topic with which the article is concerned. In the actual exposition of the articles, every effort has been made to pay close attention to the interrelationships amongst various fields of mathematics. Even when there was no need to revise the original article, changes were made in the bibliography to list items that would be more accessible to the readers.

(4) On the Appendices: The appendices have been revised and augmented along with the text. Numerical tables were deleted whenever it was felt that scientific calculators or microcomputers could easily reproduce their contents.

(5) On the Indexes: The listing of mathematical terms in Japanese has been changed, as in the case of the arrangement of articles, from the alphabetical order to the order of the Japanese syllabary. However, terms in Japanese that start with a European word have been separated out, and, in order to facilitate finding, have been listed independently in an Index for Mathematical Terms in Japanese with European Headings. The location of a term listed in the indexes is indicated by the article number and the letter heading of the section in the article where the term appears.

We here outline briefly how the process of compilation of this edition was organized. In the summer of 1978 the Mathematical Society of Japan decided, following a proposal by Professors Shōkichi Iyanaga and Yukiyo Kawada, the chief editors of the previous editions of *Sūgaku Ziten*, to undertake the compilation of the third edition. Following a resolution by the members of the Society, Kunihiro Kodaira, Sigeru Mizohata, Itiro Tamura, Nagayoshi Iwahori, Tosihisa Kimura, and myself have joined Iyanaga and Kawada to form an editorial committee and have formulated basic guidelines for the compilation of the third edition. The work of editing and compiling articles in each subject area was delegated to the editors designated for the areas, and I was asked to undertake the responsibility of putting the entire volume in order.

The subject areas and their editors are as follows:

- I Logic and Foundations: Shôji Maehara
- II Sets, General Topology, and Categories: Shôji Maehara, Hikosaburo Komatsu, Masayoshi Nagata
- III Algebra: Masayoshi Nagata
- IV Group Theory: Nagayoshi Iwahori
- V Number Theory: Takayoshi Mitsui
- VI Euclidean and Projective Geometry: Itiro Tamura
- VII Differential Geometry: Morio Obata, Shigeru Iitaka, Itiro Tamura
- VIII Algebraic Geometry: Shigeru Iitaka
- IX Topology: Itiro Tamura
- X Analysis: Sigeru Mizohata, Kiyosi Itô
- XI Complex Analysis: Yukio Kusunoki, Shigeru Iitaka
- XII Functional Analysis: Hikosaburo Komatsu
- XIII Differential, Integral, and Functional Equations: Tosihusa Kimura, Sigeru Mizohata
- XIV Special Functions: Sin Hitotumatu
- XV Numerical Analysis: Masaya Yamaguti
- XVI Computer Science and Combinatorics: Sin Hitotumatu
- XVII Probability Theory: Kiyosi Itô
- XVIII Statistics: Kei Takeuchi
- XIX Mathematical Programming and Operations Research: Shigeru Furuya
- XX Mechanics and Theoretical Physics: Huzihiro Araki
- XXI History of Mathematics: Shôkichi Iyanaga

Appendices: Sin Hitotumatu

Before the actual process of compilation started, the editorial committee had met several times, and two meetings of all the subject-area editors had been called. Furthermore, editors for each area consulted with other experts in their fields as they selected the titles of articles to be included. In areas where remarkable development had taken place, such as differential geometry, functional analysis, probability theory, and applied mathematics, extensive revision and insertion of many new articles were proposed by the respective editors. For this reason, the original articles were thoroughly reorganized and systematized. We resolved to keep the whole work in one volume, even though a number of new articles have had to be added.

The selection of articles and of their respective authors was finally completed in the spring of 1980, and 197 colleagues were asked to write the articles. The names of these contributors and those of the previous editions are listed elsewhere in this volume. To all of the authors we express our sincere gratitude.

Editing of the entire manuscript has been carried out since the summer of 1982, and we are deeply indebted to the following colleagues for their painstaking efforts in checking cross

Preface to the Third Japanese Edition

references and in proofreading: Nobuyosi Motohashi (Foundations, Set Theory), Takeo Yokonuma (Algebra, Group Theory), Takayoshi Mitsui (Number Theory), Tetsuro Kawasaki (Geometry, Topology), Shigeru Iitaka, Isao Wakabayashi (Algebraic Geometry), Morio Obata, Koichi Ogiue, Osamu Kobayashi (Differential Geometry), Seizô Ito, Hisashi Okamoto (Analysis), Hiroaki Aikawa, Makoto Ohtsuka (Computer Analysis), Hikosaburo Komatsu, Akihiko Miyachi (Functional Analysis), Kazuo Okamoto, Daisuke Fujiwara (Differential Equations), Sin Hitotumatu (Special Functions), Teruo Ushijima (Numerical Analysis), Hideo Wada (Computer Science), Yasunori Okabe (Probability Theory), Mituaki Huzii, Yoshihiro Yajima (Statistics), Shigeru Furuya (Mathematical Programming), Koichi Nakamura (Theoretical Physics), Shûichi Okamoto (History of Mathematics), Kosaku Okutsu (History of Mathematics, Number Theory).

On those occasions when it became necessary to rapidly revise manuscripts in order to unify the presentation, we were forced to go through the revision without consulting the authors of the manuscripts. I am responsible for all such revisions and hereby express my apologies to the authors concerned.

As for the indexes, we received assistance from Takeo Yokonuma, Koichi Yano, Hiroaki Aikawa, and Hisashi Okamoto in all aspects of the work, which lasted for two years, through the final proofreading. Preliminary work on the Name Index was carried out by staff members of the University of Tokyo, and Seizô Ito supervised its final compilation. We also asked many other colleagues, and in particular Nobuyuki Ikeda and Huzihiro Araki, to participate in the preparation of the Name Index, using resources available at different universities. To all of these people goes our sincere gratitude for their assistance.

Last August, when the completion of the work was drawing near, it became necessary for me to leave Japan for some time, and this made it imperative that I entrust Shigeru Iitaka with responsibility for supervising the work at the final stage of editing. I am most grateful to Iitaka for having agreed to take over this responsibility and for having brought the work to completion under the pressure of time.

From the beginning of this project, we have received an unlimited amount of assistance from the members of the Dictionary Department of Iwanami Shoten, Publishers. Messrs. Ikutaro Sasaki, Hiroto Ushida, Kazuhiko Uetake, and Nagao Sato in particular, have made supreme efforts and have come up fre-

quently with innovative ideas to make this dictionary polished and perfect. To them and also to those at Dai-Nippon Printing Co., who typeset with the use of computers the entire text, including the complicated mathematical formulas, and who have cooperated with us fully in the process of editing the indexes, goes our deep gratitude.

Kiyosi Itô
October 1985

Preface to the First English Edition

The first and second editions of *Iwanami Sūgaku Ziten* (in Japanese) were published, respectively, in April 1954 and June 1968 by Iwanami Shoten, Publishers, Tokyo. Beginning in the late 1950s, a number of unsuccessful attempts were made to arrange for translating the *Sūgaku Ziten* into European languages. Finally an agreement for an English translation was made between The MIT Press and the Mathematical Society of Japan in July 1968. The discussions were carried on first by Professor Kōsaku Yosida, then president of the Mathematical Society of Japan, and later by Yukiyo Kawada, who succeeded him in April 1968. Throughout these initial negotiations, which lasted from 1966 to 1968, we received the kindest assistance from Dr. Gordon Walker, Executive Director of the American Mathematical Society, and from Professors W. T. Martin and Shizuo Kakutani.

The agreement for the project was shortly followed by the establishment of a committee for the English edition of *Sūgaku Ziten* within the Mathematical Society of Japan, with the following membership: Professors Yasuo Akizuki, Shigeru Furuya, Sin Hito-tumatu, Masuo Hukuhara, Isao Imai, Shōkichi Iyanaga, Yukiyo Kawada, Kunihiro Kodaira, Atuo Komatu, Hirokichi Kudō, Shōji Maehara, Yukio Mimura, Kiyoshi Noshiro, Shigeo Sasaki, Shoji Ura, Nobuo Yoneda, and Kōsaku Yosida. This committee requested the original authors of the articles and other members of the Society to translate the work. A list of translators will be found at the end of this work.

In November 1968, an advisory committee for the project was formed with the following membership: Professor Edwin Hewitt (chairman), Dr. Sydney H. Gould, Professor Shizuo Kakutani, Professor Kenneth O. May, and Professor Isaac Namioka.

As the translating began, we were immediately faced with problems concerning unification of terminology and style, some of which were inherent in the differences between the structures of our two languages—for example, the fact that the Japanese language makes no distinction between singular and plural forms of nouns.

In August 1969, Professor Hewitt kindly arranged a meeting at the University of Washington, Seattle, that included the members of the Japanese and American committees and a representative of The MIT Press. It was agreed during this meeting that the transla-

tion should be faithful, with only a minimum number of changes, such as the correction of mathematical errors; whereas the references to each article might be augmented considerably for the convenience of Western readers. Professor Kenneth O. May volunteered to review the entire translated manuscript, and Professors Isaac Namioka and Shizuo Kakutani, who are proficient in both Japanese and English, proposed to read through some of the manuscript of the translated articles. It was also agreed that the Systematic List of Articles should appear in French, German, and Russian, as well as in English.

We owe very much to the American committee: Professor Hewitt organized the whole work, and Professor May revised the entire manuscript and gave us important advice concerning the appendices, according to which we deleted some of the numerical tables which may be easily found in readily accessible Western books. Professor Namioka reviewed a great part of the manuscript, transmitting his views to Professor May, who forwarded them to us with his comments. All of this assistance helped us greatly in making our final decisions. Professor Kakutani gave us very detailed and important advice on the choice of reference works.

We were also assisted concerning English terminology and reference books by the following Japanese mathematicians working in American universities: Tadatoshi Akiba, Professors Kiyosi Itô, Tatsuji Kambayashi, Tosio Kato, Teruhisa Matsusaka, Katsumi Nomizu, Ichiro Satake, Michio Suzuki, and Gaisi Takeuti.

The Mathematical Society of Japan established the following double reviewing system: group A with its twenty subgroups, each headed by the members listed below, reviewed their respective subjects; while group B reviewed the whole manuscript, mainly from the linguistic standpoint.

Group A

1. Foundations of mathematics: Shôji Maehara
2. Set theory: Atuo Komatu and Shôji Maehara
3. Algebra: Akira Hattori, Masayoshi Nagata, and Hideyuki Matsumura
4. Group theory: Shingo Murakami, Mitsuo Sugiura, and Reiji Takahashi
5. Number theory: Yuki-yosi Kawada and Tomio Kubota
6. Geometry, 7. Differential geometry: Shôkichi Iyanaga, Shigeo Sasaki, and Kentaro Yano
8. Algebraic geometry: Yasuo Akizuki and Kunihiko Kodaira
9. Topology: Atuo Komatu
10. Real analysis: Sin Hitotumatu, Shunji Kametani, and Shigeo Yano

11. Complex analysis: Kiyoshi Noshiro
12. Functional analysis: Yukio Mimura and Kôzaku Yosida
13. Differential equations: Masuo Hukuhara and Sigeru Mizohata
14. Special functions: Sin Hitotumatu
15. Numerical analysis, 16. Probability theory: Kiyosi Itô
17. Statistics: Hirokichi Kudô
18. Information theory: Tosio Kitagawa and Hirofumi Uzawa
19. Theoretical physics: Isao Imai and Kazuhiko Nishijima
20. History of mathematics: Tamotsu Murata Group B

Kenichi Iyanaga and Mitsuyo Iyanaga

Professor Sin Hitotumatu also assisted us in translating the titles of Japanese books given in the references and the explanations attached to the lists of formulas and numerical tables in the appendices. We are also grateful for the generous cooperation offered to us by our colleagues in the Department of Mathematics, Faculty of Science, University of Tokyo: Hiroshi Fujita, Shigeru Furuya, Akio Hattori, Seizô Itô, Nagayoshi Iwahori, Tosihusa Kimura, Kunihiko Kodaira, Iikosaburo Komatsu, Akihiro Nozaki, and Itiro Tamura. We are indebted as well to Professors Walter L. Bailey of the University of Chicago and Yuji Ito of Brown University for many valuable consultations concerning both mathematical and linguistic questions.

In translating the Systematic List of Articles into French, German, and Russian, we were assisted by Professor Hideya Matsumoto in Paris, Professor Emanuel Sperner in Hamburg, Professor Katsuhiro Chiba in Tokyo, and Professor Arkadii Maltsev in Moscow.

We began to send the manuscript to The MIT Press in March 1970 and finished sending it in July 1972. The manuscript was edited there, then sent to Professors May, Namioka, and Kakutani, and finally was sent back to us with their comments and questions. All of the references were carefully checked by Laura Platt.

Iwanami Shoten, Publishers, have always been cooperative with us. In the office of the Mathematical Society of Japan, Yôko Endo, Reiko Nagase, and Chieko Sagawa helped us with their efficient secretarial work.

The fruition of this project was made possible only by the gracious assistance offered to us by many people, including those already mentioned. We should like to express our most sincere gratitude to all those who have helped us so kindly.

Shôkichi Iyanaga, Yuki-yosi Kawada
August 1973

Addition made in January 1976

After the procedures described above, the whole manuscript of this Encyclopedia was sent to The MIT Press in August 1973. It was, however, toward the end of November 1975 that the final decision was made by The MIT Press to send the manuscript to composition in early 1976 in order that the work be published in 1977.

At the same time, we were asked to review and update the manuscript up to the end of February 1976. We are now making our best effort to this effect with the kind help, especially from the linguistic viewpoint, of Dr. E. J. Brody.

In so doing, we have noticed that perhaps too much emphasis has been given to results obtained by Japanese mathematicians and that there are still many things in this book which should be improved.

We hope that ongoing revisions will be carried out in subsequent editions.

Preface to the Second Japanese Edition

Seven and a half years have passed since the revised and augmented edition of *Iwanami Sūgaku Ziten* was published. The nature and purpose of this book remain the same as described in the preface to the first edition: It is an encyclopedic dictionary with articles of medium length aimed at presenting the whole of mathematics in a lucid system, giving exact definitions of important terms in both pure and applied mathematics, and describing the present state of research in each field, together with historical background and some perspectives for the future. However, mathematical science is in rapid motion, and the "present state of research" changes constantly. The present updated second edition has been published to remedy this situation as far as possible.

The main points of revision are as follows:

(1) On the Articles and the Size of the Encyclopedia: From the articles of the last edition, we have removed those whose importance has diminished recently (e.g., Geometry of Triangles), while we have added new articles in domains of growing importance (e.g., Categories and Functors; K-Theory). Many articles concerning applied mathematics in the first edition were short; in this edition, they have been combined into articles of medium length to save space and to systematize the presentation. The number of articles, 593 in the first edition, has thus diminished to 436. We have made every effort to keep the size of the encyclopedia as it was, but the substantial augmentation of content has necessarily brought about an enlargement of about 30%.

(2) On the Text: When the title of an article has remained the same as in the first edition, we have reviewed the whole text and revised whenever necessary. Especially for the fundamental ideas, we have endeavored to give thorough explanations. In the first edition we gave English, French, and German translations of article titles; in the present edition, we have also added Russian. The bibliographies at the ends of articles have been updated.

(3) On Terminology: In previous editions we endeavored to unify the terminology of the whole encyclopedia so that the reader would have no difficulties with cross-references. Here we have done this once again with the hope of attaining results more perfect than before.

(4) On the Appendices: The appendices were designed to supplement the text efficiently. Some overlapping of the appendices with the text found in previous editions has been removed. Also deleted in this edition are elemen-

tary formulas in analytic geometry and tables of statistical distributions, which can be easily found in other books. However, we have added some formulas in topology, the theory of probability, and statistics, as well as tables of characters of finite groups, etc.

(5) On the Indexes: Important terms are listed multiply in the indexes to facilitate finding, e.g., the term *transcendental singularity* appears under both *transcendental* and *singularity*. Both names in the text and those in the references are included in the Name Index of this edition. The numbers of items in the Index of Mathematical Terms in Japanese and in European Languages and in the Name Index in this edition are 17740, 10124, and 2438, compared with 8254, 8070, and 1279, respectively, in the previous edition.

The compilation of this edition was organized as follows. In the spring of 1964 we began to select the titles of articles with the aid of the following colleagues: in set theory and foundations of mathematics, Shôji Machara; in algebra and number theory, Yasuo Akizuki, Yukiyo Kawada; in differential geometry, theory of Lie groups, and topology, Yozo Matsushima, Atuo Komatu; in analysis, Masuo Hukuhara, Kôzaku Yosida, Shunji Kametani, Sin Hitotumatu; in probability theory, statistics, and mathematics for programming, Kiyosi Itô, Hirokichi Kudô, and Shigeru Furuya; in theoretical physics, Isao Imai; and for the appendices, Sin Hitotumatu. I have participated in compiling the articles on geometry and the history of mathematics. Kawada and Hitotumatu undertook the responsibility of putting the volume in order.

The work of selecting titles was completed in the summer of 1964. We then asked 173 colleagues to contribute articles. The names of these contributors and those of the previous editions are listed elsewhere. To all of them goes our most sincere gratitude.

In editing the manuscript, we were assisted by the following colleagues in addition to those mentioned already: in set theory and foundations of mathematics, Setsuya Seki and Tsurane Iwamura; in algebra and number theory, Masayosi Nagata, Akira Hattori, Hideyuki Matsumura, Ichiro Satake, Tikao Tatzawa; in geometry, theory of Lie groups, and topology, Singo Murakami, Hideki Ozeki, Noboru Tanaka, Kiiti Morita, Hiroshi Toda, Minoru Nakaoka, Masahiro Sugawara, Shôrô Araki; in analysis, Kiyoshi Noshiro, Yûsaku Komatu, Seizo Ito, Hiroshi Fujita, Shige-Toshi Kuroda, Sigeru Mizohata, Masaya Yamaguti, Tosiya Saito, Tosihusa Kimura, Masahiro Iwano; in probability theory, statistics, and mathematics of programming,

Nobuyuki Ikeda, Tadashi Ueno, Masashi Okamoto, Haruki Morimoto, Kei Takeuchi, Goro Ishii, Tokitake Kusama, Hukukane Nikaido, Toshio Kitagawa; and in theoretical physics, Ryogo Kubo, Hironari Miyazawa, Yoshihide Kozai.

After the summer of 1965, we entered into the period of finer technical editing, in which we were assisted by the following colleagues: in algebra, Keijiro Yamazaki, Shin-ichiro Ihara, Takeshi Kondo; in geometry, Tadashi Nagano, Mitsuo Sugiura, Ichiro Tamura, Kiyoshi Katase; in analysis, Nobuyuki Suita, Kotaro Oikawa, Kenkiti Kasahara, Tosinobu Muramatsu, Hikosaburo Komatsu, Setuzô Yosida, Hiroshi Tanaka; and in history, Tamotsu Murata.

We were also assisted by Katsuhiko Chiba in the Russian-language translation of the article titles, and by Osamu Kôta and Kiyoshi Katase in the indexing.

Proof sheets began appearing in the spring of 1966. In proofreading, Kaoru Sekino, Osamu Kôta, Kiyoshi Katase, and Teruo Ushijima helped us, as well as Mrs. Rieko Fujisaki. Miss Yoko Endo worked with all of us throughout the entire revision at the office of the Mathematical Society of Japan. She helped us especially in looking for and checking references and preparing the Name Index.

Yukiyo Kawada supervised the whole work, succeeding me in the role I had played in the compilation of the first and augmented editions. Sin Hitotumatu collaborated with him throughout, especially on the appendices. The second and the third proof sheets of the text were read by Kawada; the fourth proof by Hitotumatu; the proof sheets of appendices by both Kawada and Hitotumatu.

The Editorial Committee of the Mathematical Society of Japan asked me to write this preface. Having edited the previous editions and realizing full well the difficulty of the task, I would like to express my particular gratitude to Kawada. In the Dictionary Department of Iwanami Shoten, Publishers, Messrs. Hiroshi Horie, Tetsuo Misaka, Shigeaki Kobayashi, and Toshio Kouda were very cooperative in their collaboration with us. To them and also to those who typeset and printed this book at Dai-Nippon Printing Co. and Shaken Co. goes our gratitude.

S. Iyanaga
March 1968

Preface to the Revised and Augmented Japanese Edition

Six years have passed since the publication of the first edition of this encyclopedia. This revised and augmented edition incorporates the achievement of these years. It contains, together with the correction of errors found in the first edition, some new articles such as Abelian Varieties, Automata, Sheaves, Homological Algebra, Information Theory, and also supplements to articles in the first edition such as Complex Multiplication, Computers, and Manifolds. These additional items render the new edition 93 pages longer than the previous one. Each article has been thoroughly revised, and the indexes have been completely rewritten.

We were assisted by the following colleagues in selecting articles, writing, and proofreading: In set theory and foundations of mathematics, Sigeatsu Kuroda, Setsuya Seki; in algebra and number theory, Tadao Tannaka, Tsuneo Tamagawa; in real analysis, Shunji Kametani, Kôzoku Yosida; in function theory, Kiyoshi Noshiro, Shin Hitotumatu; in theory of differential and functional equations, Masuo Hukuhara, Masahiro Iwano, Ken Yamanaka; in functional analysis, Kôzoku Yosida, Seizô Itô; in geometry, Shigeo Sasaki, Nagayoshi Iwahori; in topology, Atuo Komatu, Itiro Tamura, Nobuo Yoneda; in theory of probability, Kiyosi Itô, Seizô Itô; in statistics, Toshio Kitagawa, Sigeiti Moriguti, Tatsuo Kawata; in applied mathematics, Sigeiti Moriguti; and in mechanics and theoretical physics, Takahiko Yamanouchi, Isao Imai. The revision and augmentation of the articles concerning the history of mathematics was done by myself. Miles. Yôko Tao, Eiko Miyagawa, and Mutsuko Nogami worked in the office of the Society.

The names of authors who contributed to the completion of this edition have been added to the original list of contributors.

The project of editing this edition started in the summer of 1958. We acknowledge our deep gratitude to all those who have collaborated with us since that time.

S. Iyanaga
August 1960

Preface to the First Japanese Edition

This encyclopedia, *Iwanami Sûgaku Ziten*, was compiled by the Mathematical Society of Japan at the request of the Iwanami Shoten, Publishers, who have hitherto published a series of scientific dictionaries such as *Iwanami Rikagaku-Ziten* (*Iwanami Encyclopedia of Physics and Chemistry*) and *Iwanami Tetugaku-Ziten* (*Iwanami Encyclopedia of Philosophy*). As mentioned in the prefaces to these volumes, the importance of such encyclopedias in clarifying the present state of each science is obvious if we observe the rapid pace of contemporary research. Mathematics is also in rapid motion. As a fundamental part of exact science, it serves as a basis of all science and technology. It also retains its close contact with philosophy. Therefore, the significance of having an encyclopedia of mathematics cannot be overemphasized.

Mathematics have made remarkable progress in the 20th century. As for the situation toward the end of the 19th century, we quote the following passage from the article Mathematics in the 19th Century of this encyclopedia: "Toward the end of [the 19th] century, the subjects of mathematical research became highly differentiated. Branches split into more specialized areas of study, while unexpected relations were found between previously unconnected fields. The situation became so complicated that it was difficult to view mathematics as a whole. It was in these circumstances that in 1898, at the suggestion of Franz Meyer and under the sponsorship of the Academies of Göttingen, Berlin, and Vienna, a project was initiated to compile an encyclopedia of the mathematical sciences. Entitled the *Enzyklopädie der mathematischen Wissenschaften*, it was completed in 20 years...."

One of the characteristics of 20th-century mathematics is the conscious utilization of the axiomatic method and of general concepts such as sets and mappings, which serve as foundations of different theories. Indeed, mathematics is being reorganized on the basis of topology and algebra. One such example of reorganization is found in Bourbaki's *Éléments de mathématique*; some fifteen volumes of this series have been published since 1939, and more are coming. This encyclopedia, with its limited size, cannot contain proofs for theorems. However, we intend to present a lucid view of the totality of mathematics, including its historical background and future possibilities.

Each article of this encyclopedia is of medium length—sufficiently short to permit

the reader to find exact definitions of notions, and sufficiently long to contain explanations clarifying how important concepts in the same field are related to each other. The problem of choosing adequate titles required some deliberation. The Systematic List of Articles, classified according to specific fields, shows those we have chosen. The Index of Terms contains detailed references for each notion. The appendices, including formulas and tables, supplement the text, and will be particularly useful for applied mathematicians.

The project of compiling this encyclopedia was proposed in the spring of 1947 by the Steering Committee of the Mathematical Society of Japan. It was promptly adopted, and the selection of articles in specific fields was started by the sectional committees of the Society. After seven years, our encyclopedia is finally appearing. We shall not give a detailed description of how our work proceeded through all these years. We list simply the names of those who assisted us greatly and to whom we should like to express our deep gratitude.

The president of the Society at the start of this project was the late Professor Tadahiko Kubota; but our work has been supported also by Professors Teiji Takagi, Zyoiti Suetuna, and Masatsugu Tsuji as well as by other leading members of the Society.

At the stage of selecting articles, we were assisted by the following colleagues: in history and the foundations of mathematics, Sigekatu Kuroda, Motokiti Kondo; in algebra and number theory, Kenjiro Shoda, Tadasi Nakayama, Masao Sugawara, Yukiyoji Kawada, Kenkichi Iwasawa; in geometry, Kentaro Yano, Asajiro Ichida; in function theory, Kiyoshi Noshiro, Yûsaku Komatu; in the theory of differential and functional equations, Masuo Hukuhara, Shigeru Furuya; in topology, Atuo Komatu, Ryoji Shizuma; in functional analysis, Yukio Mimura, Shizuo Kakutani, Kôzaku Yosida; in the theory of probability and statistics, Tatsuo Kawata, Toshio Kitagawa, Junjiro Ogawa; and in applied mathematics, Ayao Amemiya, Isao Imai, Kunihiko Kodaira, Shigeiti Moriguti.

We asked 190 colleagues to contribute articles, which were collected in 1949. Since then we have spent an unexpectedly long time editing them. Terminology had to be unified throughout the encyclopedia so that the reader would have no trouble with cross-references. Repetitions had to be eliminated and gaps filled. Part of the manuscript thus had to be rewritten a number of times. We have made out utmost effort in this editing work, but we are not completely without apprehension that our result has still left something to be desired.

For any shortcomings in the work, I take complete responsibility, as I have acted as the editor-in-chief. Also, since we have rewritten the manuscript, as already mentioned, we have refrained from printing the name of the original author of each article; for this, I must request the understanding of the contributors.

In the stage of editing and proofreading, we were assisted by the following colleagues: Yukio Mimura, Yukiyoji Kawada, Kazuo Matsuzaka, Sin Hitotumatu, Setsuo Fukutomi, Setsuya Seki, Shoji Irie, Shigeo Sasaki, Tatsuo Kawata, Sigekatu Kuroda, Yûsaku Komatu, Ayao Amemiya, Isao Imai, Tosio Kato, Tsurane Iwamura, Morikuni Goto, Kôzaku Yosida, Jirô Tamura, Yasuo Akizuki, Kiyoshi Noshiro, Motosaburo Masuyama, Sigeiti Moriguti, Osamu Kôta, Nobuo Yoneda, Tsuneo Tamagawa, Jun-ichi Hano; and more particularly, in the foundations of mathematics, Sigekatu Kuroda, Tsurane Iwamura; in algebra and number theory, Kazuo Matsuzaka, Yukiyoji Kawada; in algebraic geometry, Yakuo Akizuki; in real analysis, Tatsuo Kawata; in complex analysis, Yûsaku Komatu, Sin Hitotumatu, Jirô Tamura; in functional analysis, Kôzaku Yosida; in topology, Setsuo Fukutomi, Nobuo Yoneda; in the theory of probability and statistics, Motosaburo Masuyama, Sigeiti Moriguti; and in applied mathematics, Ayao Amemiya, Isao Imai, Tosio Kato, Sigeiti Moriguti.

The portraits of Abel and Riemann were kindly loaned to us by Torataro Shimomura.

The formulas in the appendices were compiled by Isao Imai, Sin Hitotumatu, and Sigeiti Moriguti; the Subject Index (in Japanese and European languages) by Osamu Kôta and Mrs. Hiroko Ide; the Name Index and Comments on Journals and Serials by Setsuo Fukutomi. Setsuo Fukutomi has taken an active part in our work ever since 1948 and given much effort to collecting and rewriting the manuscript and to unifying terminology. The editorial staff of Iwanami Shoten, Publishers, has always been cooperative. Without their generous support, this encyclopedia could never have been published.

I should like to express my sincere gratitude to all those who have collaborated with us directly or indirectly.

S. Iyanaga
March 1954

Introduction

The text of this Encyclopedic Dictionary consists of 450 *articles* arranged alphabetically, beginning with 1 Abel and ending with 450 Zeta Functions. Most of these articles are divided into *sections*, indicated by A, B, C, ..., AA, BB, Cross-references to articles, e.g., to the second article, are of the form: (\rightarrow 2 Abelian Groups) or (\rightarrow 2 Abelian Groups A), according as the whole article or a specific section is being referred to. Citations in the indexes are also given in terms of article and section numbers.

Key terms accompanied by their definitions in the text are printed in boldface. All of these terms are found in the Subject Index at the end of the volume.

The sign [†] means that the term preceded by it can be found in the index. A list of special notations used throughout the work (with explanations of their meanings) appears after the appendices.

A Systematic List of Articles, showing the general structure of the work, will be found on p. 1857 (vol. IV). The number in parentheses after each article title refers to this systematic classification into subject areas; e.g., "Abelian Varieties (VIII.5)" means that the article on Abelian varieties is the fifth article in Section VIII of the systematic list.

Books and articles in journals are cited in the text by numbers in brackets: [1], [2], At the end of each article there is a section of references in which, for books, the name of the author or authors, title, name of the publisher, year of publication, and the number of the edition are given; for journals, the name of the author, title of the article, name of the journal, and the volume numbers and inclusive page numbers are given in this order. (The names of journals and publishers are abbreviated as indicated in the lists at the end of the work.)

The Cyrillic alphabet is transliterated as follows:

Cyrillic Alphabet	Transliteration		
А	а	(a)	a
Б	б	(be)	b
В	в	(ve)	v
Г	г	(ge)	g
Д	д	(de)	d
Е	е	(ye)	e
Ё	ё	(yo)	ë
Ж	ж	(zhe)	zh
З	з	(ze)	z
И	и	(i)	i
Й	й	(i kratkoe)	ĩ
К	к	(ka)	k
Л	л	(el')	l
М	м	(em)	m
Н	н	(en)	n
О	о	(o)	o
П	п	(pe)	p
Р	р	(er)	r
С	с	(es)	s
Т	т	(te)	t
У	у	(u)	u
Ф	ф	(ef)	f
Х	х	(kha)	kh
Ц	ц	(tse)	ts
Ч	ч	(che)	ch
Ш	ш	(sha)	sh
Щ	щ	(shcha)	shch
Ъ	ъ	(tvërdyi znak)	"
Ы	ы	(yery)	y
(Ь)	ь	(myagkiĩ znak)	'
Э	э	(e)	è
Ю	ю	(yu)	yu
Я	я	(ya)	ya

A

1 (XXI.12) Abel, Niels Henrik

Niels Henrik Abel (August 5, 1802–April 6, 1829) was born a son of a poor pastor in the hamlet of Findö in Norway. In 1822, he entered the University of Christiania; however, he studied mathematics almost entirely on his own. He was recognized as a promising student by his senior, Holmboe, and after graduation he studied abroad in Berlin and Paris. In Berlin he met and was aided by A. Crelle, the founder of the *Journal für die Reine und Angewandte Mathematik*, and participated in the founding of this journal. Although he did brilliant work in Paris, he did not gain the fame he deserved. He returned to Norway in May 1827, but, unable to find a job, he was obliged to fight poverty while continuing his research. He died at twenty-six of tuberculosis.

His best-known works are the result that algebraic equations of order five or above cannot generally be solved algebraically, the result that \dagger Abelian equations can be solved algebraically, the theory of \dagger binomial series and of series in general, the theory of \dagger elliptic functions and more generally of \dagger algebraic functions, the introduction of \dagger Abelian integrals, and the establishment of \dagger Abel's theorem. His work in both algebra and analysis, written in a style conducive to easy comprehension, reached the highest level of attainment of his time.

References

- [1] N. H. Abel, *Oeuvres complètes* I, II, L. Sylow and S. Lie (eds.), Grondahl & Son, new edition, 1881 (Johnson, n.d.).
- [2] C. A. Bjerknes, Niels Henrik Abel, Gauthier-Villars, 1885.
- [3] F. Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert* I, Springer, 1926 (Chelsea, 1956).

2 (IV.2) Abelian Groups

A. General Remarks

A \dagger group G is called an **Abelian group** (or **commutative group**) if G satisfies the commutative law $ab = ba$ for all $a, b \in G$. In this article, G always denotes an Abelian group. Every \dagger subgroup of G is a \dagger normal subgroup, and all the elements of finite order in G form a subgroup

T , for which the \dagger factor group G/T has no elements of finite order except the identity e . T is called the (**maximal**) **torsion subgroup** of G . If $G = T$, then the Abelian group G is called a **torsion group** (or **periodic group**). On the other hand, if $T = \{e\}$, then G is called **torsion-free**; if $T \neq G$, $T \neq \{e\}$, then G is called **mixed**. If the order of every element of a torsion group G is a power of a fixed prime number p , then G is called an **Abelian p -group** (or **primary Abelian group**). An Abelian torsion group is the \dagger direct sum of primary Abelian groups. Thus the study of torsion groups is reduced to that of primary Abelian groups.

B. Finite Abelian Groups

The following fundamental theorem on finite Abelian groups was established by L. Kronecker, G. Frobenius, and L. Stickelberger in the 1870s. An Abelian group G of order p^n , where p is a prime number, is a direct product of \dagger cyclic subgroups Z_1, \dots, Z_r : $G = Z_1 \times \dots \times Z_r$. If Z_i is of order p^{n_i} , then $n = n_1 + \dots + n_r$, and we can assume that $n_i \geq n_{i+1}$. A direct product decomposition of G , as above, is not unique, but n_1, \dots, n_r are determined uniquely by G . The system $\{p^{n_1}, \dots, p^{n_r}\}$ or $\{n_1, \dots, n_r\}$ is called the system of **invariants** (or **type**) of G , and a system of \dagger generators $\{z_1, \dots, z_r\}$ of Z_1, \dots, Z_r is called a **basis** of G . An Abelian group of type (p, p, \dots, p) is called an **elementary Abelian group**. The decomposition of a finite Abelian group into a \dagger direct sum of subsets (not necessarily of subgroups) was considered by G. Hajós (1942) and applied successfully to a problem of number theory (\rightarrow 151 Finite Groups).

C. Finitely Generated Abelian Groups

The theory of \dagger finitely generated Abelian groups, i.e., Abelian groups generated by a finite number of elements, is as old as that of finite Abelian groups. The direct product of \dagger infinite cyclic groups is called a **free Abelian group**. A finitely generated Abelian group G is the direct product of a finite Abelian group and a free Abelian group. The finite factor is the **torsion subgroup** of G . The free factor of the group G is not necessarily unique; however, the number of infinite cyclic factors of the free factor is uniquely determined and is called the **rank** of G . Two finitely generated Abelian groups are isomorphic if they have isomorphic maximal torsion subgroups and the same rank. This theory can be extended to the theory of \dagger modules over a \dagger principal ideal domain (\rightarrow 67 Commutative Rings K).

D. Torsion Groups

The structure of Abelian p -groups is relatively well known, compared with other infinitely generated Abelian torsion groups. In the 1920s, H. Prüfer made the first important contribution to the study of Abelian p -groups, and H. Ulm and L. Zippin completed the theory for countable groups in the 1930s. The uncountable case was first treated by L. Kulikov in the 1940s, but the study of this case is still in progress.

An Abelian p -group $G \neq \{e\}$ is called **divisible** (or **complete**) if for any $a \in G$ there is an element $x \in G$ satisfying $x^p = a$. A divisible group is a \dagger direct sum of **Abelian groups of type p^∞** (Prüfer). Here a group of type p^∞ is isomorphic to the \dagger multiplicative group of all the p^n th roots of unity ($n = 1, 2, \dots$) in the complex number field. Let G be any Abelian p -group. The maximal divisible subgroup V of G is a direct product factor of G : $G = V \times R$, where R has no divisible subgroups. An Abelian p -group without a divisible subgroup is called a **reduced Abelian group**.

An element x of an Abelian p -group G is said to have **infinite height** if for any n there is an element $y_n \in G$ satisfying $x = y_n^{p^n}$. The elements of infinite height form a subgroup G^1 of G . If $G^1 = \{e\}$ and G is countable, then G is decomposed uniquely into the direct sum of cyclic groups. This assertion fails if the hypothesis of countability is dropped. By \dagger transfinite induction we can define G^β as follows. If β is an \dagger isolated ordinal number, then $G^\beta = (G^{\beta-1})^1$; if β is a \dagger limit ordinal number, then $G^\beta = \bigcap_{\alpha < \beta} G^\alpha$. For the least ordinal number τ such that $G^\tau = G^{\tau+1}$, G^τ is the maximal divisible subgroup of G . If G is reduced, then $G^\tau = \{e\}$. We call τ the **type** of an Abelian p -group G . For $\alpha < \tau$, $\bar{G}^\alpha = G^\alpha / G^{\alpha+1}$ is called the **Ulm factor** of G , and the sequence $\bar{G}^0, \dots, \bar{G}^\alpha, \dots$ ($\alpha < \tau$) is called the **sequence of Ulm factors** of G . Each Ulm factor \bar{G}^α has no element of infinite height, and if $\alpha < \tau - 1$, then \bar{G}^α has an element of arbitrarily large order. Let τ be a countable ordinal number; assume that for any ordinal number $\alpha < \tau$ there is given a countable Abelian p -group A_α such that A_α has no element of infinite height, and that for $\alpha \neq \tau - 1$ A_α has an element of arbitrarily large order. Then there is a reduced countable Abelian p -group which is of type τ with a sequence of Ulm factors isomorphic to $A_0, A_1, \dots, A_\alpha, \dots$ ($\alpha < \tau$) (Zippin). Two reduced countable Abelian p -groups A and B are isomorphic if they have the same type τ ; and for any $\alpha < \tau$ the Ulm factors \bar{A}^α and \bar{B}^α are isomorphic. The assertion fails if the hypothesis of countability is dropped.

E. Torsion-Free Groups

In Abelian groups, the group operation is often written $a + b$, using the additive notation; an additively written group, called an **additive group**, is generally assumed to be Abelian. In the rest of this article we consider exclusively additive Abelian groups, of which the additive group \mathbf{Z} of rational integers is the most primitive example. In such a group the identity element is called the **zero element** and is denoted by 0; the inverse of a is denoted by $-a$, and we write $a + (-b) = a - b$. The direct sum of additive groups A_λ ($\lambda \in \Lambda$) is called a **free additive group** if each A_λ is isomorphic to \mathbf{Z} . An additive group G is regarded as a \dagger module over the \dagger ring \mathbf{Z} , to which the notion of linear independence is applicable (\rightarrow 277 Modules). Elements a_1, \dots, a_r of G are **linearly dependent** if there are integers n_1, \dots, n_r , not all of which are zero such that $n_1 a_1 + \dots + n_r a_r = 0$. Those that are not linearly dependent are termed **linearly independent**. An infinite set of elements of G is called linearly independent if the elements of any finite subset are linearly independent. If there are N elements of G that are linearly independent, but if any $N + 1$ elements of G are linearly dependent, then N is called the **rank** of G . Such a system of N linearly independent elements is called a **maximal independent system**. A torsion-free additive group G is not necessarily free if G is not finitely generated.

The first important work on torsion-free additive groups was done by F. W. Levi (1917). A. G. Kurosh (1937) completed the theory in the case of finite rank. In the general case little is known, and I. Kaplansky, J. Rotman, and others are continuing the investigation.

The additive group \mathbf{Q} of rational numbers is of rank 1, and conversely any additive group of rank 1 is isomorphic to some subgroup of \mathbf{Q} . An additive group G is called **divisible** (or **complete**) if for any $a \in G$ and for any integer n there is an element $x_n \in G$ such that $nx_n = a$. A divisible torsion-free additive group is isomorphic to a direct sum of some copies of \mathbf{Q} . For any torsion-free additive group G there is a divisible torsion-free additive group containing G . A minimal additive group F among these groups is uniquely determined up to isomorphism and has the rational number field \mathbf{Q} as an \dagger operator domain. Let $\mathbf{Q}^{(p)} = \{a/b \mid (a, b) = 1, p \nmid b\}$ be the ring of p -integers in \mathbf{Q} , and let G_p be the smallest $\mathbf{Q}^{(p)}$ -subgroup of F containing G . Let \mathbf{Q}_p be the $\dagger p$ -adic number field and \mathbf{Z}_p the ring of $\dagger p$ -adic integers. Extending the operator domain \mathbf{Q} to \mathbf{Q}_p we obtain naturally a \mathbf{Q}_p -module F_p from F . Let \bar{G}_p be the natural closure of G_p in F_p . Then \bar{G}_p has \mathbf{Z}_p as an

operator domain and thus becomes a \mathbb{Z}_p -module. A \mathbb{Z}_p -module of rank N is isomorphic to the direct sum of κ_p copies of \mathbb{Q}_p and $N - \kappa_p$ copies of \mathbb{Z}_p : $G_p \cong \sum_n \mathbb{Q}_p v_n + \sum_m \mathbb{Z}_p w_m$ ($n = 1, \dots, \kappa_p$; $m = 1, \dots, N - \kappa_p$). Here κ_p is called the p -rank of G . As the invariants of G , Kurosh gives the rank, the p -ranks for all primes p , and a certain equivalence class of the sequence of the matrices \mathfrak{M}_p . Here p ranges over all primes, and \mathfrak{M}_p is the matrix of coefficients when the elements of a maximal independent system of F are written as linear combinations of (v_n, w_m) .

F. General Abelian Groups

An Abelian group is, in general, an extension of a torsion group by a torsion-free group. A torsion group T is called **bounded** if there is an integer n such that $t^n = 1$ for all $t \in T$. Suppose there is a torsion group T . Then T is a direct summand of an Abelian group G which contains T as its maximal torsion subgroup if and only if T is the direct product of a divisible group and a bounded group (R. Baer and S. Fomin).

G. Characters

A **character** χ of an Abelian group G is a function which assigns to each $a \in G$ a complex number $\chi(a)$ of absolute value 1 and satisfies $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in G$. The product $\chi = \chi_1 \chi_2$ of two characters χ_1 and χ_2 is defined by $\chi(a) = \chi_1(a)\chi_2(a)$, and χ is also a character of G . Thus all the characters of G form an Abelian group $C(G)$, which is called the **character group** of G . The identity element of the character group is the **identity character** (or **principal character**) χ such that $\chi(a) = 1$ for all $a \in G$. If G is finite, then $G \cong C(G)$. This implies the duality $G = C[C(G)]$. This fact was extended by L. S. Pontryagin to locally compact topological Abelian groups (\rightarrow 422 Topological Abelian Groups B–D). For additive Abelian groups with operator domains \rightarrow 277 Modules.

References

- [1] L. Fuchs, Abelian groups, Pergamon, 1960.
- [2] I. Kaplansky, Infinite Abelian groups, Univ. of Michigan, 1954.
- [3] A. G. Kurosh, Theory of groups I, II, Chelsea, 1960. (Original in Russian, 1953.)

3 (VIII.5) Abelian Varieties

A. History

Except for C. F. Gauss, whose work on this subject saw the light many years after his death, N. H. Abel was the first to consider algebraic functions as functions of complex variables and to discover double periods of elliptic functions such as $x = x(u)$, which is the inverse function of an elliptic integral

$$u = \int^x \frac{dx}{\sqrt{f_4(x)}},$$

where $f_4(x)$ is a polynomial in x of degree 4. C. G. Jacobi expressed elliptic functions explicitly as ratios of theta series. As a natural generalization of elliptic functions, which are Abelian functions of genus 1, Jacobi and his successors studied the inverse function of the hyperelliptic integral, or more generally, the inverse function of the Abelian integral. By investigating hyperelliptic integrals of the first kind of genus 2,

$$\int^{x_1} \frac{dx_1}{\sqrt{f_6(x_1)}}, \quad \int^{x_2} \frac{x_2 dx_2}{\sqrt{f_6(x_2)}},$$

where $f_6(x)$ is a polynomial of degree 6 in x , Jacobi obtained multiple-valued functions with quadruple periods. He then discovered the remarkable fact that if we consider sums of two integrals

$$u_1 = \int^{x_1} \frac{dx_1}{\sqrt{f_6(x_1)}} + \int^{x_2} \frac{dx_2}{\sqrt{f_6(x_2)}},$$

$$u_2 = \int^{x_1} \frac{x_1 dx_1}{\sqrt{f_6(x_1)}} + \int^{x_2} \frac{x_2 dx_2}{\sqrt{f_6(x_2)}},$$

the elementary symmetric functions $s_1 = x_1 + x_2$ and $s_2 = x_1 x_2$ of x_1 and x_2 are single-valued functions of u_1 and u_2 with quadruple periods. He also conjectured that these functions s_1 and s_2 might be expressed explicitly in terms of theta series of u_1 and u_2 ; this conjecture was later confirmed by J. G. Rosenhain and A. Göpel.

In the latter half of the 19th century, the general theory of Abelian functions was established. The central subject was Jacobi's inverse problem (\rightarrow Section L), which is a natural generalization of the above results. B. Riemann solved this problem by expressing elementary Abelian functions (\rightarrow Section M) as rational functions of theta functions.

The theory of functions with multiple periods was developed further by H. Poincaré, G. Frobenius, and E. Picard. In the 20th cen-

tury, the importance of the theories of Abelian functions and Abelian varieties has become more obvious with the development of the theory of functions of several complex variables and algebraic geometry. In particular, problems intimately related to number theory have given rise to the development of a purely algebraic theory of Abelian varieties. Valuable contributions have been made by S. Lefschetz, C. L. Siegel (*Siegel modular forms), A. Weil (algebraic theory of Abelian varieties, systematization of the theory of theta functions), and by D. Mumford (*moduli theory, algebraic theory of theta functions).

B. Algebraic Theory

When a \dagger group variety is \dagger complete as a variety, the group law is commutative; such a group variety is called an **Abelian variety** (\rightarrow 12 Algebraic Varieties H.). Let B be a subvariety of an Abelian variety A , and assume that B is a subgroup of A as an abstract group. Then B has the structure of an Abelian variety whose law of composition is induced by that of A , and B is called an **Abelian subvariety** of A . More generally, when an algebraic subset \mathfrak{B} is a subgroup of A , then the component B of \mathfrak{B} containing the identity element is an Abelian subvariety, and \mathfrak{B} is a union of B and a finite number of cosets of B . When A is defined over a field k , then any Abelian subvariety of A is defined over a finite \dagger separable extension of k (W. L. Chow's theorem). An Abelian variety A is called **simple** if A has no Abelian subvarieties other than A itself and 0.

Every \dagger rational mapping of an algebraic variety V into an Abelian variety is defined at each simple point of V . This implies that an Abelian variety is \dagger absolutely minimal.

C. Homomorphisms

A rational mapping of an Abelian variety A into an Abelian variety B is called a **rational homomorphism** (or simply **homomorphism**) if f is a group homomorphism. Let F be a rational mapping of A into B ; then F can be uniquely expressed as follows: $F(x) = F_0(x) + F(0)$ ($x \in A$), where F_0 is a homomorphism and $F(0)$ is the image of the unit element 0 of A . Hence the structure of an Abelian variety (as a group variety) is uniquely determined by the underlying algebraic-variety structure.

When a rational homomorphism f is birational, f is called a **birational isomorphism** (or simply **isomorphism**). It is clear that a rational isomorphism is an abstract isomorphism, but the converse is not necessarily true. Let A, B

be two Abelian varieties. We say that A is **isogenous** to B if the dimension of A is equal to that of B and there exists a surjective homomorphism of A onto B , or equivalently, if there exists a surjective homomorphism of A onto B whose kernel is finite. The relation of isogeny is an equivalence relation. For an Abelian variety A and an Abelian subvariety X of A there is an Abelian subvariety Y of A such that the natural homomorphism $X \times Y \rightarrow A: (x, y) \rightarrow x + y$ is an isogeny (**Poincaré's complete reducibility theorem**). In particular, every Abelian variety is isogenous to a product of simple Abelian varieties that are determined uniquely up to isogeny and order.

Let A, B be two Abelian varieties; we denote by $\text{Hom}(A, B)$ the additive group of rational homomorphisms of A into B . When a rational homomorphism λ is surjective, then the **degree** $v(\lambda)$ of λ is defined by $\lambda(A) = v(\lambda)B$ as \dagger algebraic cycles. If λ is an isogeny, then $v(\lambda) \neq 0$, and the order of the kernel $\{t \mid t \in A, \lambda t = 0\}$ is at most $v(\lambda)$ and equal to $v(\lambda)$ if and only if λ is \dagger separable. The additive group $\text{Hom}(A, B)$ is shown to be free of rank $\leq 4 \dim A \cdot \dim B$ by the l -adic representation (\rightarrow Section E). If $A = B$, then $\text{Hom}(A, A)$ has a ring structure; it is called the **ring of endomorphisms** (or **endomorphism ring**) of A and is denoted by $\mathfrak{U}(A)$. The tensor product $\mathfrak{U}_0(A) = \mathfrak{U}(A) \otimes \mathbf{Q}$, where \mathbf{Q} is the field of rational numbers, is an \dagger associative algebra over \mathbf{Q} . If A is simple, then $\mathfrak{U}_0(A)$ is a \dagger division algebra. More generally, $\mathfrak{U}_0(A)$ is isomorphic to a direct product of some \dagger total matrix algebras over division algebras; thus $\mathfrak{U}_0(A)$ is \dagger semisimple. In particular, if A is 1-dimensional (in other words, if A is an \dagger elliptic curve), the types of $\mathfrak{U}_0(A)$ are well known; when the characteristic $p = 0$, then $\mathfrak{U}_0(A)$ is either the field of rational numbers or an \dagger imaginary quadratic field. When $p > 0$, aside from these two fields, we have a \dagger quaternion algebra over \mathbf{Q} as a possible type of $\mathfrak{U}_0(A)$.

Let k be a finite field with q elements. An algebraic integer is called a **Weil number** for q if every conjugate of it has absolute value $\sqrt[q]{q}$. If A is an Abelian variety defined and simple over k , the q th power endomorphism of $A: x \rightarrow x^q$ determines a conjugacy class of Weil numbers for q , as Weil showed (\rightarrow 450 Zeta Functions). Moreover, we have the following classification theorem (J. Tate, T. Honda): There is a one-to-one correspondence between the set of all k -isogeny classes of k -simple Abelian varieties over k and the set of all conjugacy classes of Weil numbers for q . Tate also determined the structure of the division algebra $\mathfrak{U}_0(A)$ over \mathbf{Q} , which is described in terms of the decomposition of the q th power endomorphism into prime ideals.

D. Divisors

Let \mathfrak{G} be the additive group of \dagger divisors on an Abelian variety A and \mathfrak{G}_a be the subgroup of divisors that are \dagger algebraically equivalent to 0. Then the factor group $\mathfrak{G}/\mathfrak{G}_a$ has no \dagger torsion part; this implies that for an Abelian variety \dagger numerical equivalence coincides with \dagger algebraic equivalence. We denote this relation by \equiv . Given an element a of A , the translation $T_a: A \ni x \rightarrow x+a \in A$ gives a \dagger birational transformation, which is everywhere \dagger biregular, on the underlying variety of A ; we denote by X_a the image of a divisor X on A . Then $X \equiv 0$ if and only if X_a is \dagger linearly equivalent to X for each point a of A . The \dagger Albanese variety of an Abelian variety A is A itself, and the \dagger Picard variety \hat{A} of A is isogenous to A . In particular, for the \dagger Jacobian variety J of an algebraic curve, \hat{J} is isomorphic to J itself. The Picard variety $\hat{\hat{A}}$ of \hat{A} is isomorphic to A (**duality theorem**). Let X be a divisor on A ; the mapping $a \rightarrow$ the \dagger linear equivalence class of the divisor $X_a - X$, $a \in A$, is a rational homomorphism of A into \hat{A} , and we denote it by φ_X . If $\varphi_X = 0$, then $X \equiv 0$, and vice versa; hence we have an additive map of $\mathfrak{G}/\mathfrak{G}_a$ into $\text{Hom}(A, \hat{A})$. If φ_X is surjective, we say that X is **nondegenerate**. A \dagger positive divisor X is nondegenerate if and only if X is \dagger ample, and then nX is \dagger very ample for $n \geq 3$. There always exist positive nondegenerate divisors on an Abelian variety; therefore an Abelian variety is a \dagger projective variety. For a given divisor X on A , we can find n suitable points u_1, \dots, u_n , where n is the dimension of A , so that the \dagger intersection product $X_{u_1} \cdot \dots \cdot X_{u_n}$ is defined. We denote by $(X^{(n)})$ the \dagger degree of the zero cycle $X_{u_1} \cdot \dots \cdot X_{u_n}$. If X is positive nondegenerate, then the dimension $l(X)$ of the \dagger defining module of the \dagger complete linear system determined by X is equal to $(X^{(n)})/n!$ (**Poincaré's theorem**). Furthermore, the degree $v(\varphi_X)$ of φ_X , where X is any divisor on A , is given by the formula $v(\varphi_X) = ((X^{(n)})/n!)^2$ (**Frobenius's theorem**).

For a nondegenerate divisor X on A there is a unique integer $i = i(X)$, $0 \leq i \leq \dim A$, called the **index** of X , such that $H^p(A, \mathcal{O}(X)) = 0$ for $p \neq i$ and $H^i(A, \mathcal{O}(X)) \neq 0$. We have $i(-X) = \dim A - i(X)$ and $i(X) = 0$ if and only if X is ample.

E. l -adic Representations

Let A be an Abelian variety of dimension n . For a given prime number l , let $\mathfrak{G}_l(A)$ denote the group of points on A whose order is a power of l . If l is different from the characteristic of the base field of A , then the group $\mathfrak{G}_l(A)$ is isomorphic to the direct product of $2n$

factor groups $\mathbf{Q}_l/\mathbf{Z}_l$, where \mathbf{Q}_l is the field of l -adic numbers and \mathbf{Z}_l is the group of l -adic integers (\rightarrow 439 Valuations). We call such an isomorphism the **l -adic coordinate system** of $\mathfrak{G}_l(A)$. Now let λ be a rational homomorphism of A into an Abelian variety B of dimension m . Then we can see that λ induces a homomorphism of $\mathfrak{G}_l(A)$ into $\mathfrak{G}_l(B)$. This shows that by placing l -adic coordinate systems on $\mathfrak{G}_l(A)$ and $\mathfrak{G}_l(B)$ respectively, we get a matrix representation $M_l(\lambda)$ of λ with $2m$ rows and $2n$ columns. The representation $\lambda \rightarrow M_l(\lambda)$ is faithful, and $M_l(\lambda)$ is called the **l -adic representation** of λ . When $A = B$, then $\lambda \rightarrow M_l(\lambda)$ is a faithful representation of the ring of endomorphisms $\mathfrak{A}(A)$. This representation can be naturally extended to the representation of the algebra $\mathfrak{A}_0(A)$; the characteristic polynomial of the l -adic representation $M_l(\lambda)$ (where λ is an element of $\mathfrak{A}_0(A)$) is a polynomial with coefficients in \mathbf{Q} . Moreover, the polynomial does not depend on the choice of the prime number l . When $\lambda \in \mathfrak{A}(A)$, then $v(\lambda)$ is equal to $\det M_l(\lambda)$. The trace of $M_l(\lambda)$ is usually written as $\sigma(\lambda)$.

Let λ be a rational homomorphism of A into B and Y be a divisor on B . Then by the correspondence $\text{cl}(Y) \rightarrow \text{cl}(\lambda^{-1}(Y))$, we obtain a rational homomorphism of \hat{B} into \hat{A} , called the **transpose** of λ and denoted by ${}^t\lambda$, where cl means the linear equivalence class (\rightarrow 16 Algebraic Varieties M). If X is a nondegenerate divisor on A , then the composition map $\beta: \hat{A} \rightarrow A$ of ${}^t\varphi_X: \hat{A} \rightarrow \hat{\hat{A}}$ and the canonical isomorphism $\hat{\hat{A}} \rightarrow A$ satisfy the equality $\beta \circ \varphi_X = v(\varphi_X)\delta$ ($\delta =$ the identity mapping of A). We denote by φ_X^{-1} the element $(1/v(\varphi_X))\beta$ in $\text{Hom}(\hat{A}, A) \otimes \mathbf{Q}$. The correspondence $\alpha \rightarrow \alpha'$, $\alpha' = \varphi_X^{-1} \circ {}^t\alpha \circ \varphi_X$ ($\alpha \in \mathfrak{A}_0(A)$) is an \dagger involution of $\mathfrak{A}_0(A)$ and is of order 1 or 2. If $\alpha \neq 0$, then $\sigma(\alpha' \circ \alpha) > 0$ (**Castelnuovo's lemma**). A. Weil was the first to recognize the importance of this theorem in connection with \dagger Riemann's hypothesis on \dagger congruence zeta functions.

F. Differential Forms

A \dagger differential form ω on an Abelian variety A of dimension n is called **invariant** if $T_a^*\omega = \omega$ for every point $a \in A$, where T_a is the translation by a (\rightarrow Section D). The differential form \dagger of the first kind is invariant, and conversely, every invariant differential form is of the first kind. Let K be the \dagger universal domain and $K(A)$ be the \dagger function field of A . The set of the linear differential forms of the first kind on A is a linear space over K of dimension n , and its basis becomes a basis over $K(A)$ of the linear space consisting of all linear differential forms on A . An **invariant derivation** on A is a derivation D in $K(A)$ satisfying $(Df) \circ T_a = D(f \circ T_a)$

for any element f of $K(A)$ and every point a of A . For a linear differential form $\omega = \sum f_i du_i$ and a derivation D , we put $\langle \omega, D \rangle = \sum f_i Du_i$. Then $\langle \omega, D \rangle$ is a bilinear form in ω and D . A derivation D is invariant if and only if $\langle \omega, D \rangle$ is a constant function for every invariant linear differential form ω . Similarly, a linear differential form ω is invariant if and only if $\langle \omega, D \rangle$ is a constant function for every invariant derivation D . The linear space of invariant linear differential forms and that of invariant derivations are dual to each other with respect to the bilinear form $\langle \omega, D \rangle$.

Now consider the case when the characteristic p of the universal domain is positive. The automorphism $a \rightarrow a^p$, $a \in K$, of the universal domain K induces a group isomorphism of A ; we denote by A^p the image of A and by x^p the image of a point x of A . The image A^p is an Abelian variety, and the group isomorphism $\pi: x \rightarrow x^p$, $x \in A$, is an isogeny of A onto A^p . Let B be another Abelian variety and let λ be an isogeny of A onto B . If there is an isogeny $\mu: B \rightarrow A^p$ such that $\pi = \mu \circ \lambda$, then we say that λ is of **height 1**. The function field $K(B)$ of B can be considered as a subfield of the function field $K(A)$ of A by the mapping λ . If λ is of height 1 and $v(\lambda) = p$, there exists an invariant derivation D of $K(A)$ with the constant field $K(B)$, uniquely determined up to constants. Moreover, we can choose D so that $D^p = D$ or $D^p = 0$. In the first case λ is said to be of type (i_1) ; in the second case it is said to be of type (i_2) . An isogeny whose degree is a prime different from the characteristic p is said to be of type (s_1) , and a separable isogeny whose degree is p is said to be of type (s_2) . Any isogeny can be written as a product of isogenies of these four types.

G. Polarized Abelian Varieties

Let X be a divisor on an Abelian variety A ; we denote by \mathfrak{X} the class of divisors X' such that $mX \equiv m'X'$ for suitably chosen positive integers m, m' . When the class \mathfrak{X} contains positive nondegenerate divisors, we say that \mathfrak{X} determines a **polarization** on A , and the couple (A, \mathfrak{X}) is called a **polarized Abelian variety**. In particular, if A is a \dagger Jacobian variety whose polarization \mathfrak{X} is determined by a theta divisor, we call (A, \mathfrak{X}) the **canonically polarized Jacobian variety**. If an endomorphism α of A keeps the polarization invariant, i.e., if the class determined by $\alpha^{-1}(X)$ coincides with the class \mathfrak{X} , then α is called an **endomorphism** of the polarized Abelian variety (A, \mathfrak{X}) . In particular, if α is an automorphism of A , then we say that α is an **automorphism** of (A, \mathfrak{X}) . The group of all automorphisms of a polarized

Abelian variety is finite. In particular, the group of automorphisms of a canonically polarized Jacobian variety is finite. Hence follows the famous theorem concerning the finiteness of the group of automorphisms of an \dagger algebraic curve of genus not less than 2.

On the other hand, the algebraic equivalence class of a nondegenerate divisor is called an **inhomogeneous polarization**. (The above polarization is then sometimes called a homogeneous polarization.) An inhomogeneous polarization X determines an isogeny $\varphi_X: A \rightarrow \hat{A}$ uniquely. If φ_X is an isomorphism, the polarization is called **principal**. An endomorphism of an inhomogeneously polarized Abelian variety can be defined similarly.

H. Analytic Theory

For the rest of this article we take the complex number field \mathbb{C} as the universal domain, and in this case we can utilize analytic and topological methods.

Let \mathbb{C}^n be an n -dimensional vector space over \mathbb{C} . In a natural way, the space \mathbb{C}^n becomes a $2n$ -dimensional vector space \mathbb{R}^{2n} over the real field \mathbb{R} , and the mapping $J: z \rightarrow \sqrt{-1}z$, $z \in \mathbb{C}^n$, is an \mathbb{R} -linear automorphism of \mathbb{R}^{2n} such that $J^2 = -1$. Conversely, if for an even-dimensional \mathbb{R} -vector space \mathbb{R}^{2n} such a mapping J is given, then by putting $(a + \sqrt{-1}b)x = ax + bJx$ ($x \in \mathbb{R}^{2n}$; $a, b \in \mathbb{R}$), we can introduce an n -dimensional complex linear structure into \mathbb{R}^{2n} . We then say that J determines a **complex structure** on \mathbb{R}^{2n} ; we denote by $\mathbb{C}^n = (\mathbb{R}^{2n}, J)$ the space having the complex structure determined by J . Let $\omega_1, \dots, \omega_{2n}$ be $2n$ \mathbb{R} linearly independent points on $\mathbb{C}^n = (\mathbb{R}^{2n}, J)$. Then the \dagger lattice \mathbb{D} generated by these points is discrete, and the factor group $\mathbb{T}^n = \mathbb{C}^n / \mathbb{D}$ is a **complex torus** of dimension n . We fix a basis of \mathbb{C}^n and introduce a complex coordinate system on \mathbb{C}^n . Utilizing the basis $\omega_1, \dots, \omega_{2n}$ of \mathbb{R}^{2n} , we also introduce a real coordinate system on \mathbb{R}^{2n} . We then obtain an $n \times 2n$ matrix $\Omega = (\omega_{ij})$, where the $(\omega_{1i}, \dots, \omega_{ni})$ are the complex coordinates of ω_i ; the matrix Ω is called the **period matrix** of \mathbb{T}^n . Let (z_1, \dots, z_n) be the complex coordinates of a point $z \in \mathbb{C}^n$ and (x_1, \dots, x_{2n}) be the real coordinates. Then we have $\dagger(z_1, \dots, z_n) = \Omega \dagger(x_1, \dots, x_{2n})$. If we let the same symbol J stand for the representation matrix of the linear transformation J with respect to the basis $\omega_1, \dots, \omega_{2n}$, then we have $\sqrt{-1}\Omega = \Omega J$. The underlying real Lie group of \mathbb{T}^n is a $2n$ -dimensional (real) \dagger torus group $(\mathbb{R}/\mathbb{Z})^{2n}$. Hence the \dagger Poincaré polynomial of \mathbb{T}^n is $(1+x)^{2n}$. Any \dagger Hermitian metric on \mathbb{C}^n (as a vector space) induces a \dagger flat \dagger Kähler metric; hence the

Hodge numbers (\rightarrow Kähler Manifolds C) are $h^{p,q} = \binom{n}{p} \binom{n}{q}$.

I. Theta Functions

A holomorphic function $f(z)$ on $C^n = (R^n, J)$ is called a **theta function** if for every $d \in D$ we have $f(z+d) = f(z) \exp(l_d(z) + c_d)$, where $l_d(z)$ is a linear form on C^n which, as for c_d , depends on d . The set of zeros of a theta function f , which we write as (f) , determines an analytic divisor on T^n . Conversely, for every effective analytic divisor X on T^n , there exists a theta function f such that $(f) = X$. With respect to the real coordinate system x_1, \dots, x_{2n} determined by the basis $\omega_1, \dots, \omega_{2n}$, we can find $2n \times 2n$ matrices A, A_0 and a $1 \times 2n$ matrix b , with elements in C , such that the **transformation formula** $f(x+a) = f(x) \exp(2\pi\sqrt{-1}(aAx + \frac{1}{2}aA_0a + ba))$ (where $A \equiv A_0 \pmod{Z}$, $A_0 = A_0$) holds for every $1 \times 2n$ matrix a whose elements are rational integers. Moreover, if we put $E = A - {}^1A$, then E is an alternating matrix whose elements are rational integers, and $S = EJ$ is a positive semidefinite symmetric matrix. Conversely, if there exists such an alternating matrix E we can find a theta function. (There does exist, however, a complex torus on which no theta function exists other than trivial ones, i.e., ones of the form $\exp(\varphi(z))$, where $\varphi(z)$ is a polynomial of degree at most 2.)

A theta function f is called **nondegenerate** if it cannot be a function of $n-1$ complex variables, and f is nondegenerate if and only if the matrix $S = EJ$ is positive definite. A complex torus has the structure of an Abelian variety if and only if there exists a nondegenerate theta function, i.e., if and only if there exists an alternating matrix E whose elements are rational integers such that EJ is a positive definite symmetric matrix. The latter condition is satisfied if and only if there exists an alternating matrix E whose elements are rational integers such that $\Omega {}^1E^{-1} \Omega = 0$, $\sqrt{-1} \Omega {}^1E^{-1} \Omega > 0$ (positive definite Hermitian matrix). In particular, a period matrix Ω satisfying these conditions is called a **Riemann matrix**, and the rational matrix ${}^1E^{-1}$ is called the **principal matrix** belonging to Ω .

Determining a polarization on an Abelian variety can be reduced to designating a class of principal matrices obtained from a given

from f is invertible. For a given principal matrix ${}^1E^{-1}$, we can choose suitable coordinate systems of C^n and D so that $\Omega = (I_n, F)$ and $E = \begin{pmatrix} 0 & \Delta \\ -\Delta & 0 \end{pmatrix}$, where I_n is a unit matrix and Δ is a diagonal matrix whose elements are elementary divisors of E . In these situations, ΔF is symmetric and its imaginary part positive definite, i.e., ΔF is a point of the Siegel upper half-space \mathfrak{S}_n . Thus to the polarized Abelian variety (where the polarization is determined by ${}^1E^{-1}$) there corresponds a point of \mathfrak{S}_n . This gives a one-to-one correspondence between the isomorphism classes of Abelian varieties inhomogeneously polarized by a principal matrix with given elementary divisors and the points of the factor space $\mathfrak{S}_n/\Gamma_n(\Delta)$, where $\Gamma_n(\Delta)$ is a subgroup that is commensurable to the Siegel modular group of degree n and operates on \mathfrak{S}_n discontinuously. The case of principal polarization corresponds to the case when $\Delta = I_n$, and then $\Gamma_n(\Delta)$ is the Siegel modular group itself. $\mathfrak{S}_n/\Gamma_n(\Delta)$ is the coarse moduli space of Abelian varieties, polarized as above; the projective embedding of $\mathfrak{S}_n/\Gamma_n(\Delta)$ is given by means of Siegel modular forms.

J. Abelian Functions

Meromorphic periodic functions on C^n with periods $\omega_1, \dots, \omega_{2n}$ (i.e., meromorphic functions on $T = C^n/D$, $D = Z\omega_1 \oplus \dots \oplus Z\omega_{2n}$) are called **Abelian functions** with periods D (or on T). The quotient of two theta functions having the same periods and the same transformation formulas is an Abelian function, and conversely any Abelian function can be written as such. All Abelian functions on T form a field $C(T)$ called the **Abelian function field**. If T has the structure of an Abelian variety, then $C(T)$ coincides with the field of rational functions on T . In general, for any complex torus T there are an Abelian variety T' (possibly of dim 0) and a surjective homomorphism $\lambda: T \rightarrow T'$ which induces an isomorphism $\lambda^*: C(T') \simeq C(T)$. Such a pair $(T', \lambda: T \rightarrow T')$ is unique up to isomorphism, and is called the **algebraic reduction** of T .

K. Homomorphisms

Let $C^{n_i} = (R^{2n_i}, J_i)$ ($i = 1, 2$) be complex linear spaces, D_1, D_2 be lattices in R^{2n_1}, R^{2n_2} respectively.

$T^n = C^n/D$, and let f be a theta function such that $(f) = X$. Then the divisor X is nondegenerate if and only if the theta function f is nondegenerate, and the latter statement holds if and only if the alternating matrix E obtained

$i = 1, 2$. If a C -linear mapping $\Lambda: C^{n_1} \rightarrow C^{n_2}$ satisfies $\Lambda(D_1) \subset D_2$, then Λ induces a complex analytic homomorphism of $T_1 = C^{n_1}/D_1$ to $T_2 = C^{n_2}/D_2$. Conversely, any complex analytic homomorphism of T_1 to T_2 is ob-

tained in this way. Let \mathbf{T}_1 and \mathbf{T}_2 be Abelian varieties, and let $\Omega_1 = (\omega_1^{(1)}, \dots, \omega_{2n_1}^{(1)})$ and $\Omega_2 = (\omega_1^{(2)}, \dots, \omega_{2n_2}^{(2)})$ be their Riemann matrices. Then for a homomorphism $\lambda: \mathbf{T}_1 \rightarrow \mathbf{T}_2$, we can find a representation matrix $W(\lambda)$ with complex coefficients; and with respect to the real coordinate systems $(\omega_1^{(1)}, \dots, \omega_{2n_1}^{(1)})$ and $(\omega_1^{(2)}, \dots, \omega_{2n_2}^{(2)})$, we can find a representation matrix $M(\lambda)$ with coefficients in \mathbf{Z} such that $W(\lambda)\Omega_1 = \Omega_2 M(\lambda)$. Conversely, if for a complex matrix W there is a matrix M , with coefficients in \mathbf{Z} , satisfying the relation given above, then W gives a homomorphism of \mathbf{T}_1 to \mathbf{T}_2 . The above equation is called **Hurwitz's relation**. The notion of l -adic coordinate system, which is valid for a general characteristic, corresponds to that of the lattice group, and the l -adic representation $M_l(\lambda)$ of λ is the abstraction of the integral representation $M(\lambda)$.

L. Abelian Integrals

Let \mathfrak{R} be a compact \dagger Riemann surface of genus $g \geq 1$ (\rightarrow 11 Algebraic Functions) and let ω be a sum of \dagger Abelian differentials of the first kind or of the second kind. Then the \dagger period of ω along a cycle γ depends only on the \dagger homology class of γ . The set of all differentials of the first kind forms a complex linear space of dimension g ; we denote it by \mathfrak{D}_0 . Let P be a point and P_0 a fixed point of \mathfrak{R} ; then we denote by $u(P)$ the vector integral $(\int_{P_0}^P \omega_1, \dots, \int_{P_0}^P \omega_g)$, where $(\omega_1, \dots, \omega_g)$ is a basis of \mathfrak{D}_0 and the path from P_0 to P is common to every integral. The correspondence $P \rightarrow u(P)$ is not a single-valued mapping; the totality of differences of values of $u(P)$ coincides with the group \mathbf{D} consisting of periods $(\int_\gamma \omega_1, \dots, \int_\gamma \omega_g)$, where γ varies over all cycles. Let a set of cycles $\{\gamma_1, \dots, \gamma_{2g}\}$ be a basis of the homology group, with coefficients in \mathbf{Z} ; then $2g$ column vectors of the $g \times 2g$ matrix $\Omega = (\omega_{ij})$, $\omega_{ij} = \int_{\gamma_j} \omega_i$, are linearly independent over \mathbf{R} . Since the group \mathbf{D} coincides with the set of linear combinations, with coefficients in \mathbf{Z} , of column vectors of Ω , this \mathbf{D} is a \dagger lattice of rank $2g$, and the matrix Ω is a period matrix of the complex torus $\mathbf{T}^g = \mathbf{C}^g/\mathbf{D}$ of dimension g .

For a basis of the homology group with coefficients in \mathbf{Z} , we take \dagger normal sections $\alpha_1, \dots, \alpha_g, \alpha_{g+1}, \dots, \alpha_{2g}$ of \mathfrak{R} , and let the same symbol Ω stand for the period matrix (ω_{ij}) , $\omega_{ij} = \int_{\alpha_j} \omega_i$; then we have

$$\Omega E^* \Omega = 0, \quad \sqrt{-1} \Omega E^* \bar{\Omega} > 0$$

(positive definite Hermitian matrix), where

$$E = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix} \quad (\text{with } I_g \text{ the unit matrix}).$$

This implies that E is a principal matrix be-

longing to Ω ; we call the equality and inequality just given **Riemann's period relation** and **Riemann's period inequality**, respectively. Furthermore, we can choose a suitable basis of \mathfrak{D}_0 so that the period matrix Ω is of the form (I_g, F) with $F \in \mathfrak{S}_g$ (\rightarrow Section I). We consider the function $\mathfrak{g}(u)$, $u = (u_1, \dots, u_g)$, defined by an infinite series

$$\mathfrak{g}(u) = \sum_m \exp(2\pi \sqrt{-1} (\mathbf{m}^* u + \frac{1}{2} \mathbf{m}^* F \mathbf{m})),$$

where the sum is taken over all row vectors $\mathbf{m} = (m_1, \dots, m_g)$ with coefficients in \mathbf{Z} . If u is in a bounded region, then the series for $\mathfrak{g}(u)$ is uniformly and absolutely convergent. Hence $\mathfrak{g}(u)$ is a holomorphic function of u . This is a theta function corresponding to the principal matrix $E = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix}$ and is called the

Riemann theta function. As $\mathfrak{g}(u)$ is nondegenerate, the complex torus $\mathbf{T}^g = \mathbf{C}^g/\mathbf{D}$ has the structure of an Abelian variety. If we regard the Riemann surface \mathfrak{R} as an algebraic curve, this Abelian variety \mathbf{T}^g is precisely the Jacobian variety of the curve \mathfrak{R} . The collection of zeros of $\mathfrak{g}(u)$, a divisor on \mathbf{T}^g , defines the canonical polarization on the Jacobian variety (\rightarrow Section G).

The correspondence $P \rightarrow u(P)$, $p \in \mathfrak{R}$ induces a well-defined mapping φ of \mathfrak{R} into $\mathbf{T}^g = \mathbf{C}^g/\mathbf{D}$. Moreover, if we set $\varphi(A) = \sum_{j=1}^r \varphi(P_j) - \sum_{j=1}^r \varphi(Q_j)$ for any divisor $A = P_1 \dots P_r / Q_1 \dots Q_r$ of degree 0, then $\varphi(A)$ is a point on $\mathbf{T}^g = \mathbf{C}^g/\mathbf{D}$, which is represented by the vector $(\sum_{j=1}^r \int_{Q_j}^{P_j} \omega_1, \dots, \sum_{j=1}^r \int_{Q_j}^{P_j} \omega_g)$ of \mathbf{C}^g , and the mapping $A \rightarrow \varphi(A)$ is a homomorphism of the group \mathfrak{G}_0 of divisors of degree 0 onto $\mathbf{T}^g = \mathbf{C}^g/\mathbf{D}$. The kernel of this homomorphism coincides with the group \mathfrak{G}_1 of \dagger principal divisors (**Abel's theorem**). Hence a divisor $A = P_1 \dots P_r / Q_1 \dots Q_r$ of degree 0 is a divisor of some function if and only if we have $\sum_{j=1}^r \int_{Q_j}^{P_j} \omega_i \equiv 0 \pmod{\mathbf{D}}$ ($i = 1, \dots, g$), or equivalently, the left-hand side is 0 for a suitable path.

Given g fixed points P_1, \dots, P_g on \mathfrak{R} and given (u_1, \dots, u_g) as any vector of \mathbf{C}^g , the problem of finding g points Q_1, \dots, Q_g satisfying relations $\sum_{j=1}^g \int_{P_j}^{Q_j} \omega_i \equiv u_i \pmod{\mathbf{D}}$, $i = 1, \dots, g$, is called **Jacobi's inverse problem**. To solve the problem, we take a divisor A of degree 0 such that the class $\varphi(A) \pmod{\mathbf{D}}$ is represented by (u_1, \dots, u_g) ; then, by virtue of the \dagger Riemann-Roch theorem, there exists a divisor $Q_1 \dots Q_g$ satisfying

$$A \equiv Q_1 \dots Q_g / P_1 \dots P_g \pmod{\mathfrak{G}_1}.$$

Abel's theorem implies that the set of points $\{Q_1, \dots, Q_g\}$ is a solution of Jacobi's problem. Moreover, for general (u_1, \dots, u_g) , the solution is unique; i.e., there exists a subvariety \mathfrak{X} of dimension $g-2$ on \mathbf{C}^g such that the solution is

unique up to order if and only if (u_1, \dots, u_g) does not lie on $\tilde{\mathfrak{X}}$. In particular, if every point P_i coincides with the fixed point P_0 that appeared in the definition of $u(P)$, the subvariety on $\mathbf{T}^g = \mathbf{C}^g/\mathbf{D}$ determined by $\tilde{\mathfrak{X}}$ is obtained in the following way: Let $W_1 + \dots + W_{2g-2}$ be a canonical divisor on \mathfrak{R} , and put $c = \varphi(W_1) + \dots + \varphi(W_{2g-2})$. Then the locus of points $c - \varphi(R_1) - \dots - \varphi(R_{g-2})$ (where $g-2$ points R_1, \dots, R_{g-2} are taken independently over all points of \mathfrak{R}) is the desired subvariety.

In terms of the theory of complex manifolds the above result can be rewritten as follows: Let $S^g(\mathfrak{R})$ be the symmetric product of \mathfrak{R} and $\tilde{\mathfrak{X}}$ the subvariety of $\mathbf{T}^g = \mathbf{C}^g/\mathbf{D}$ induced from $\tilde{\mathfrak{X}}$. Then the holomorphic map $S^g(\varphi): S^g(\mathfrak{R}) \rightarrow \mathbf{T}^g$ induced from $\varphi: \mathfrak{R} \rightarrow \mathbf{T}^g$ is bimeromorphic and isomorphic outside that $\tilde{\mathfrak{X}}$ whose elements correspond to effective divisor classes of degree g contained in canonical divisors on \mathfrak{R} .

M. Elementary Abelian Functions

Let z be a nonconstant meromorphic function on \mathfrak{R} . Then for any $u = (u_1, \dots, u_g) \in \mathbf{C}^g$ that does not lie on $\tilde{\mathfrak{X}}$, there exist g points Q_1, \dots, Q_g (uniquely determined up to order) as the solution of Jacobi's problem. Therefore the elementary symmetric functions

$$s_1(u; z) = \sum_{j=1}^g z(Q_j),$$

$$s_2(u; z) = \sum_{i < j} z(Q_i)z(Q_j), \dots, s_g(u; z) = \prod_{j=1}^g z(Q_j)$$

are well defined if u lies outside the variety $\tilde{\mathfrak{X}}$ of dimension $g-2$. Each function $s_i(u; z)$, regarded as a function of u , can be extended uniquely to an Abelian function in the whole space \mathbf{C}^g ; the function is denoted by the same symbol $s_i(u; z)$. The Abelian functions $s_1(u; z), \dots, s_g(u; z)$ are called the **elementary Abelian functions** obtained from z .

Now let K be the field of Abelian functions on \mathbf{C}^g/\mathbf{D} and k the field of meromorphic functions on \mathfrak{R} ; then the dimension of K over \mathbf{C} is g , and $[K: \mathbf{C}(s_1(u; z), \dots, s_g(u; z))] = r^g$, where r is the degree of the function z that is given by $[k: \mathbf{C}(z)]$. Moreover, if we take any function w such that $k = \mathbf{C}(z, w)$, then we have

$$K = \mathbf{C}(s_1(u; z), \dots, s_g(u; z); s_1(u; w), \dots, s_g(u; w)),$$

where $s_i(u; w), \dots, s_g(u; w)$ are the elementary Abelian functions obtained from w .

We can write any elementary Abelian function as a rational function of Riemann theta functions; therefore any Abelian function can be written as a rational function of Riemann theta functions. Furthermore, if u and v are variable vectors, then $s_i(u+v; z)$ can be repre-

sented as an algebraic function of $s_1(u; z), \dots, s_g(u; z), s_1(v; z), \dots, s_g(v; z)$; i.e., we can choose a suitable polynomial $H_i(Z; X_1, \dots, X_g; Y_1, \dots, Y_g)$ with coefficients in \mathbf{C} so that $H_i(s_i(u+v; z); s_1(u; z), \dots, s_g(u; z); s_1(v; z), \dots, s_g(v; z)) = 0$. This **algebraic addition formula** with respect to the elementary Abelian functions $s_i(u; z)$, $i = 1, \dots, g$, is a function-theoretic interpretation of the fact that the addition map $X \times X \rightarrow X: (x, y) \rightarrow x + y$ is a t morphism of algebraic varieties.

As the study of t Abelian integrals of the first kind led us to the theory of Jacobian varieties, t Abelian integrals of the second and the third kind give rise to the theory of t generalized Jacobian varieties (\rightarrow 9 Algebraic Curves).

The theory of Abelian varieties has significant applications to number theory, as shown by the following examples: the theory of t unramified Abelian extensions with respect to a function field of several variables defined over a finite field (S. Lang), the theory of heights of points on an Abelian variety (Weil, A. Néron, J. Tate), and the theory of complex multiplication (\rightarrow 73 Complex Multiplication) in the case of higher dimensions (Y. Taniyama, G. Shimura).

N. Some Recent Results

(1) Level Structure, Moduli of Abelian

Varieties. Let A be an Abelian variety over k of dimension g and n a positive integer which the characteristic of k does not divide. A **level n structure** on A is defined to be a set of $2g$ points $\sigma_1, \dots, \sigma_{2g}$ on A which form a basis for the group of points of order n on A .

Let $A(g, d, n; k)$ be the set of triples: (i) an Abelian variety A over k of dimension g , (ii) an inhomogeneous polarization X on A with $v(\varphi_X) = d^2$, and (iii) a level n structure $\sigma_1, \dots, \sigma_{2g}$ of A , all up to isomorphisms. Similarly we can define $A(g, d, n; S)$ for Abelian schemes over a scheme S . The correspondence $S \rightarrow A(g, d, n; S)$ defines a functor $\mathcal{A}(g, d, n)$.

D. Mumford has shown that there exists the t coarse moduli scheme $A(g, d, n)$ quasi-projective over $\text{Spec}(\mathbf{Z}[1/n])$, and that it is even fine if $n \geq 3$ [6]. He used the technique of t Hilbert schemes and t stable points (\rightarrow 16 Algebraic Varieties). One of the key steps of his proof is to show that for an embedding $\phi: A \rightarrow \mathbf{P}^m$ of degree r (i.e., the degree of $\phi(A)$ in \mathbf{P}^m) over an algebraically closed field k and a positive integer n such that $\text{char}(k)$ does not divide n and $n > \sqrt{(m+1)r}$, the point $(\phi(x_i))_{i=1, \dots, n^{2g}}$ in $(\mathbf{P}^m)^{n^{2g}}$ is stable with respect to the action of $\text{PGL}(m)$, where the x_i are the points of order n on A (with an arbitrary order). Mumford

later showed another method of constructing the moduli of polarized Abelian varieties by using algebraic theta constants [7].

(2) Néron Minimal Models, Good and Stable

Reduction. Let R be a discrete valuation ring with residue field k and quotient field K . For an Abelian variety A over K , there exists a smooth group scheme \mathcal{A} of finite type over $S = \text{Spec}(R)$, called the **Néron minimal model** of A , such that for every scheme S' smooth over S there is a canonical isomorphism

$$\text{Hom}_S(S', \mathcal{A}) \cong \text{Hom}_K(S'_K, A),$$

where S'_K is the pullback of S' by $\text{Spec}(K) \rightarrow \text{Spec}(R)$ (A. Néron, M. Raynaud). In particular, we have $\mathcal{A}_K \cong A$. Denote by A_0 the fiber of \mathcal{A} over the closed point of S .

If \mathcal{A} is proper over S , we say that A has a **good reduction** at R . If the connected component A_0^0 of A_0 containing 0 has no unipotent radical (or equivalently, A_0^0 is an extension of an Abelian variety by an ℓ -algebraic torus over a finite algebraic extension of k), we say that A has a **stable reduction**. If there is a finite separable extension K' of K with a prolongation R' of R to K' such that $A \times_K K'$ has good (stable) reduction, we say that A has a **potential good (stable) reduction** at R . Let K_s be a separable closure of K and \bar{R} a prolongation of R to K_s . For a prime number $l \neq \text{char}(k)$, we have a canonical homomorphism $\rho: \text{Gal}(K_s/K) \rightarrow \text{Aut}(\mathcal{G}_l(A))$, called a monodromy. Then A has potential good reduction if and only if the image of the ℓ -inertia group $I(\bar{R})$ by ρ is a finite group (J.-P. Serre and Tate). Every Abelian variety A over K has potential stable reduction at R (**stable reduction theorem**, A. Grothendieck [16]).

(3) Graded Ring of Theta Functions. If f is a theta function with a period system \mathbf{D} , f^n is also a theta function with the same period system for every positive n . We denote by S_n the vector space of the theta functions with the period system \mathbf{D} subject to the same transformation law as f^n . If we denote by X the effective divisor on $\mathbf{T} = \mathbf{C}^g/\mathbf{D}$ defined by f , then S_n can be naturally considered as the ℓ -defining module of the ℓ -complete linear system of nX , and the dimension of $S_n (= l(nX))$ is equal to the product of the nonzero diagonal elements of $n\Delta$ (\rightarrow Section I) (Frobenius). If $g \in S_m$ and $h \in S_n$, then $gh \in S_{m+n}$; hence $S = \bigoplus_{n \geq 0} S_n$ is a ℓ -graded ring, which is normal and finitely generated. For $m \geq 2$ and $n \geq 3$, the product map $S_m \times S_n \rightarrow S_{m+n}$ is surjective (D. Mumford, S. Koizumi). If the elementary divisors of E can be divided by an integer ≥ 4 , the kernel of the natural graded mapping $S(S_1) \rightarrow S$ (where

3 Ref.

Abelian Varieties

$S(S_1)$ denotes the ℓ -symmetric algebra over S_1) is generated by the quadratic relations (i.e., the part of degree 2) for sufficiently large degrees (Mumford). Geometrically this means that if $X = (f)$, $f \in S_1$, is nondegenerate, then, with the projective embedding $i: \mathbf{T} \rightarrow \mathbf{P}^N$ defined by the complete linear system of X , $i(\mathbf{T})$ is an intersection of quadrics in \mathbf{P}^N containing $i(\mathbf{T})$.

Mumford has developed a theory of algebraic theta functions that also works for the positive characteristic case [7] and has proved the above results in general.

References

- [1] Sém. H. Cartan, vol. 2, Fonctions analytiques de plusieurs variables I–V, 1951–1952, Benjamin, 1967.
- [2] F. Conforto, Abelsche Funktionen und algebraische Geometrie, Springer, 1956.
- [3] A. Krazer and W. Wirtinger, Abelsche Funktionen und allgemeine Thetafunktionen, Enzykl. Math. Wiss., II, 2, B7 (1920), 604–873.
- [4] J. Igusa, Theta functions, Springer, 1972.
- [5] S. Lang, Abelian varieties, Interscience, 1959.
- [6] D. Mumford, Geometric invariant theory, Erg. Math., Springer, 1965; second edition, 1982.
- [7] D. Mumford, On the equations defining Abelian varieties I–III, Inventiones Math., 1 (1966), 287–354; 3 (1967), 75–135, 215–244.
- [8] D. Mumford, Varieties defined by quadratic equations, Questioni sulle varietà algebriche, Corsi dal C. I. M. E., Cremona, 1969, III Ciclo, Cremonese, 1970, 31–94.
- [9] D. Mumford, Abelian varieties, Tata Inst. studies in math., Oxford Univ. Press, 1974.
- [10] A. Néron, Modèles minimaux des variétés abéliennes sur les corps locaux et globaux, Publ. Math. Inst. HES, 21 (1964), 5–128.
- [11] W. F. Osgood, Lehrbuch der Funktionentheorie II₂, Teubner, 1932.
- [12] M. Raynaud, Modèles de Néron, C. R. Acad. Sci. Paris, 262 (1966), 345–347.
- [13] B. Riemann, Theorie der Abelschen Funktionen, J. Reine Angew. Math., 54 (1857), 115–155.
- [14] J.-P. Serre, Quelques propriétés des variétés abéliennes en caractéristique p , Amer. J. Math., 80 (1958), 715–739.
- [15] J.-P. Serre and J. Tate, Good reduction of Abelian varieties, Ann. Math., (2) 88 (1968), 492–517.
- [16] Séminaire de géométrie algébrique (SGA) 7, Groupes de monodromie en géométrie algébrique, Lecture notes in math. 288, 340, Springer, 1972, 1973.

Additive Number Theory

- [17] G. Shimura, Moduli and fibre systems of Abelian varieties, *Ann. Math.*, (2) 83 (1966), 294–338.
- [18] C. L. Siegel, Analytic functions of several complex variables, Lecture notes, Institute for Advanced Study, Princeton, 1948–1949.
- [19] J. Tate, Classes d'isogénie des variétés abéliennes sur un corps fini (d'après T. Honda), *Sém. Bourbaki*, 1968–1969, no. 352; Lecture notes in math. 179, Springer, 1971.
- [20] A. Weil, Théorèmes fondamentaux de la théorie des fonctions thêta, *Séminaire Bourbaki*, no. 16, 1949, Benjamin, 1966.
- [21] A. Weil, Variétés abéliennes et courbes algébriques, *Actualités Sci. Ind.*, Hermann, 1948.
- [22] A. Weil, Variétés kähleriennes, *Actualités Sci. Ind.*, Hermann, 1958.
- [23] A. Weil, Généralisation des fonctions abéliennes, *J. Math. Pures Appl.*, (9) 17 (1938), 47–87.
- [24] D. Mumford, Tata lecture notes on theta functions I, II, Birkhäuser, 1983.

4 (V.5) Additive Number Theory

A. General Remarks

Let \mathbf{N} be the set of natural numbers, and let A, B, \dots be subsets of \mathbf{N} . The sum $C = A + B$ is defined as the set $\{c | c \in A, c \in B, \text{ or } c = a + b, a \in A, b \in B\}$. A finite sum of subsets of \mathbf{N} is defined similarly. If \mathbf{N} is the sum $A + \dots + A$ (r times), then we say that A is a **basis of order r in \mathbf{N}** . Let $A(x)$ denote the number of integers in A that do not exceed x . The **density** of A is defined as $\inf A(x)/x$. If $A(n) \geq \alpha n$, $B(n) \geq \beta n$ ($0 < \alpha, \beta < 1$) for all $n \in \mathbf{N}$, then we have $(A+B)(n) \geq (\min(1, \alpha + \beta))n$. This result was stated by E. Landau (1937) without proof; A. J. Khinchin had given a proof (1932) for the case $\alpha = \beta$, and H. B. Mann (1942) and E. Artin and P. Scherk (1943) succeeded in proving the statement for the general case. Suppose that the densities of A and B are, respectively, α and 0. If B is a basis of finite order in \mathbf{N} , then the density of $A + B$ is greater than that of A (Khinchin and P. Erdős, 1936). Let P be the set of all prime numbers. Though the density of P is 0, the density of $P + P$ is positive (L. G. Shnirel'man, 1930). Hence P is a basis of finite order in \mathbf{N} ; in other words, there exists a positive integer r such that every natural number can be expressed as a sum of at most r primes. Though the density of the set Q of the k th powers of natural numbers is 0, there exists a positive integer $s(k)$ such that the sum $Q + \dots$

$+ Q$ ($s(k)$ times) is a set of positive density. L. K. Hua (1956) gave a simple proof of this fact, based on Yu. V. Linnik's idea. It follows, therefore, that any natural number n can be expressed as a sum $n = a_1^k + \dots + a_r^k$, where $a_i \in \mathbf{N}$, $t \geq s(k)$. This result had already been shown by D. Hilbert (1909).

An ancient method of finding prime numbers is 'Eratosthenes' sieve. V. Brun (1920) devised a new sieve method to express an arbitrary integer n as the sum of two integers $n = a + b$, where the number of prime factors of a and b is as small as possible. This method was improved by H. A. Rademacher (1924), Landau (1931), A. Bustab (1937), and others. Among these the method found by A. Selberg (1952) is notable (\rightarrow 123 Distribution of Prime Numbers E).

B. Farey Sequences

Let τ be a positive integer. We arrange in increasing order the set of all positive irreducible fractions lying between 0 and 1 whose denominators do not exceed τ . This sequence is called the **Farey sequence** of order τ . For example, the Farey sequence of order 5 consists of

$$0, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, 1.$$

A necessary and sufficient condition that a fraction a/b be directly followed by a fraction c/d in the Farey sequence of order n is $b + d \geq n + 1$, $bc - ad = 1$. In this case the fraction $(a + c)/(b + d)$ is called the **mediant** of a/b and c/d . Interpolating the Farey sequence of order n with such mediants $(a + c)/(b + d)$ satisfying $b + d = n + 1$, we obtain the Farey sequence of order $n + 1$.

Let a/q be a fraction in the Farey sequence of order τ , and $a'/q', a''/q''$ be adjacent members of a/q in the sequence such that $a'/q' < a/q < a''/q''$. The interval $[(a' + a)/(q' + q), (a + a'')/(q + q'')]$ is known as the **Farey arc** surrounding a/q . In particular, if $a/q = 0/1$, then we set $[-1/n, 1/n]$ to be the Farey arc surrounding 0/1, where $n = [\tau] + 1$ ($[\]$ is the Gauss symbol). We can thus decompose the interval $[-1/n, 1 - 1/n]$ into a disjoint union of Farey arcs. If α is contained in the Farey arc surrounding a/q , then $|\alpha - a/q| < 1/q\tau$. Therefore, for a given $\tau \geq 1$ and a real α , we can prove the existence of a/q such that

$$(|a|, q) = 1, \quad 0 < q \leq \tau, \quad |\alpha - a/q| < 1/q\tau.$$

Let $f(\zeta) = \sum_{n=0}^{\infty} a_n \zeta^n$ be a power series which is convergent in the disk $|\zeta| \leq 1$. Then, by Cauchy's integral formula,

$$a_n = \frac{1}{2\pi i} \int_{|\zeta|=1} \frac{f(\zeta)}{\zeta^{n+1}} d\zeta = \int_0^1 f(e^{2\pi i \alpha}) e^{-2\pi i n \alpha} d\alpha.$$

To estimate such an integral, we sometimes utilize the decomposition of the interval $[0, 1]$ into a disjoint union of Farey arcs as mentioned above. This method is called the **circle method**, and the subdivision of the interval is known as the **Farey dissection**.

Given a positive number c , a Farey arc around a/q is usually called a **major arc** (or **basic interval**) if q does not exceed the given bound c ; otherwise, it is called a **minor arc** (or **supplementary interval**). Usually, the principal part of the previously mentioned integral is derived from the integral over the major arcs, and the residual part is provided by the integral over the minor arcs.

C. Goldbach's Problem

Goldbach's problem is found in letters (1742) he exchanged with L. Euler. In them he stated that every positive integer can be expressed as the sum of primes. More precisely, he conjectured that any even integer not smaller than 6 can be expressed as the sum of two odd primes and that any odd integer not smaller than 9 can be expressed as the sum of three odd primes.

I. M. Vinogradov (1937) proved that every sufficiently large odd integer can be expressed as the sum of three primes. Let N be a sufficiently large odd integer. If we write

$$A(q, N) = \frac{\mu(q)}{\varphi^3(q)} \sum_{\substack{1 \leq a \leq q \\ (a, q)=1}} \exp\left(-2\pi i \frac{a}{q} N\right),$$

then the series $S(N) = \sum_{q=1}^{\infty} A(q, N)$ is absolutely convergent and is equal to

$$\prod_p \left(1 + \frac{1}{(p-1)^3}\right) \prod_{p|N} \left(1 - \frac{1}{p^2 - 3p + 3}\right).$$

It is known that $S(N) > 6/\pi^2$ for all N . If we denote by $r(N)$ the number of solutions of $N = p_1 + p_2 + p_3$, then $r(N) \sim (N^2/2(\log N)^3) \cdot S(N)$. To prove this, Vinogradov used the circle method. He employed the \dagger prime number theorem for arithmetic progressions to estimate the integrals over the major arcs and devised an ingenious method to estimate the series $\sum_{p \leq N} \exp(2\pi i \alpha p)$ in the computation of the integrals on the minor arcs. A finite or infinite sum of exponential functions such as this is called a **trigonometric sum**. More generally, we consider trigonometric sums of several variables. Vinogradov provided detailed remarks and calculations [6].

In the case of even integers, the problem is still unsolved, although J. G. van der Corput, T. Estermann, and N. G. Chudanov proved

4 D

Additive Number Theory

simultaneously (1938) that almost all even integers (i.e., except a set of density 0) can be expressed as the sum of two primes. For these problems, Linnik (1946) and Chudanov (1947) introduced function-theoretic methods. They obtained the density theorem concerning the zeros of L -series, and A. Zulauf (1952) continued along the same lines. These methods had been suggested by G. H. Hardy and J. E. Littlewood (c. 1919), although they had assumed that the \dagger extended Riemann hypothesis held.

We denote by P_j a number having at most j prime factors counted with their multiplicity. Using his sieve method, Brun (1919) proved that every sufficiently large even integer $2n$ can be expressed in the form $2n = P_9 + P_9$. Rademacher (1924) improved this result and obtained $2n = P_7 + P_7$. Applying his new sieve method to this problem, Selberg (1950) proved that $2n = P_2 + P_3$. On the other hand, using the \dagger large sieve, A. Rényi proved that $2n = P_1 + P_k$ for some k . Afterwards, combining Richert's sieve with a large sieve, Chen Jing-Run (1973) proved that $2n = P_1 + P_2$ for large $2n$ [11].

D. Polygonal Numbers

Let m be an integer greater than 3, and let $a_1 = 1$, $a_{n+1} - a_n = (m-2)n + 1$ ($n = 1, 2, \dots$). The sequence $\{a_n\}$ forms the system of **polygonal numbers of order m** . The general term of $\{a_n\}$ is given by $n + \frac{1}{2}(m-2)(n^2 - n)$ ($n = 1, 2, \dots$). Such a_n are said to be **triangular numbers** if $m = 3$, **square numbers** if $m = 4$, and **pentagonal numbers** if $m = 5$.

P. Fermat (1636) stated that every natural number can be expressed as the sum of m polygonal numbers of order m . This conjecture was proved by A. M. Legendre (1798) for $m = 3$, by J. L. Lagrange (1772) for $m = 4$, and by A. L. Cauchy (1813) for the general case. With regard to Lagrange's result, Legendre noticed that in order that a positive integer n be expressed as the sum of three squares, it is necessary and sufficient that n not be of the form $4^a(8m+7)$.

Given a positive integer n , the number of integral solutions of the equation $x_1^2 + x_2^2 + \dots + x_s^2 = n$ is denoted by $r_s(n)$. For example, $r_2(5) = 8$. The \dagger generating function $\sum_{n=1}^{\infty} r_2(n)n^{-s}$ can be expressed as $\sum_{m,n=-\infty}^{+\infty} (m^2 + n^2)^{-s}$, where the term corresponding to $m = n = 0$ is omitted; the function is equal to $4\zeta_K(s)$, where $\zeta_K(s)$ is the \dagger Dedekind zeta function of the Gaussian number field $K = \mathbf{Q}(\sqrt{-1})$. The equation $4\zeta_K(s) = 4\zeta(s)L(s, \chi)$ (where $\chi(n) = (-4/n)$) leads to

$$r_2(n) = 4 \sum'_{m|n} (-1)^{(m-1)/2},$$

where \sum' means the sum over all odd factors m of n . This result was obtained by C. G. J. Jacobi (1829). He also obtained the following formula:

$$r_4(n) = 8 \sum'_{m|n, 4 \nmid m} m,$$

where \sum' means the sum over all divisors m of n not divisible by 4 (Hardy and E. M. Wright, C. L. Siegel, 1964). Let $q = \exp(2\pi i\tau)$ ($\text{Im } \tau > 0$), and

$$f(q) = 1 + \sum_{n=1}^{\infty} r_s(n)q^n \\ = (1 + 2q + 2q^4 + 2q^9 + \dots)^s.$$

Hardy (1920) considered the variation of $f(q)$ for $q = a \exp(2\pi i h/k)$ (with $0 < a < 1$) as $a \rightarrow 1$; he obtained

$$f(q) \sim \pi^{s/2} \left(\frac{S_{h,k}}{k} \right)^s \log \left(\frac{1}{a} \right)^{-s/2},$$

where $S_{h,k} = \sum_{j=1}^k \exp(2\pi i(h/k)j^2)$. Furthermore, he constructed the **singular series**

$$\rho_s(n) = \frac{\pi^{s/2} n^{(s/2)-1}}{\Gamma(s/2)} \sum_{k=1}^{\infty} A_k, \quad A_1 = 1,$$

$$A_k = k^{-s} \sum_{\substack{1 \leq h \leq k \\ (h,k)=1}} (S_{h,k})^s \exp \left(-2\pi i \frac{h}{k} n \right),$$

and showed that $r_s(n) = \rho_s(n)$ for $s = 5, 6, 7$, and 8. P. T. Bateman (1951) proved the same equation for $s = 3$ and 4. If $s \geq 9$ then $r_s(n) = \rho_s(n) + O(n^{s/4})$ (Hardy, Littlewood, and S. Ramanujan, c. 1919). The detailed exposition of this result is in the notes of A. Z. Walfisz (1952) and Rademacher [4]. H. D. Kloostermann (1926) and Estermann (1962) studied the equation $ax_1^2 + bx_2^2 + cx_3^2 + dx_4^2 = n$, which led to a new field of study concerning the \dagger Kloostermann sum. For instance, estimates such as

$$\left| \sum_{x=1}^{p-1} \exp \left(2\pi i \frac{cx + d/x}{p} \right) \right| \leq 2\sqrt{p}$$

are obtained by using the theory of zeta functions in algebraic function fields in one variable (A. Weil, 1948) (\rightarrow 450 Zeta Functions).

E. Waring's Problem

The first formulation of **Waring's problem** is found in E. Waring, *Meditationes algebraicae* (1770), in which he discusses the problem of expressing an arbitrary positive integer as the sum of at most nine cubes or as the sum of at most nineteen biquadratics. Hilbert proved (\rightarrow Section A) that there exists a positive integer $s(k)$ such that, for any integer N , the equation

$$x_1^k + x_2^k + \dots + x_s^k = N$$

has a nonnegative integral solution if $s \geq s(k)$. We denote by $g(k)$ the least value of $s(k)$, and by $G(k)$ the least value of $s(k)$ for which the equation is solvable with at most finitely many exceptions of N . Research concerning $g(k)$ and $G(k)$ received its initial impetus from the circle method considered by Hardy and Littlewood, and it underwent considerable development in the works of H. Weyl and Vinogradov.

Let $r_s(N)$ be the number of solutions of the above equation. We then have

$$r_s(N) = \int_0^1 \left(\sum_{x \leq N^{1/k}} \exp(2\pi i \alpha x^k) \right)^s \times \exp(-2\pi i N \alpha) d\alpha.$$

If we make the Farey dissection, translating the interval $[0, 1]$ slightly, then the main term of $r_s(N)$ is provided by the integrals over major arcs, and the residual term is derived from the integrals over minor arcs. Let p be a prime, and let $M(N, p^l)$ denote the number of solutions of the congruence equation

$$x_1^k + x_2^k + \dots + x_s^k \equiv N \pmod{p^l}.$$

Then $\lim_{l \rightarrow \infty} M(N, p^l)/p^{l(s-1)} = \chi_p(N)$ is not zero, and the infinite product $\prod_p \chi_p(N) = S(N)$ converges for $s \geq 4k$, where $S(N)$ is larger than a positive constant which is determined independently of the choice of N . On the other hand, let

$$S(a, q) = \sum_{x=0}^{q-1} \exp \left(2\pi i \frac{a}{q} x^k \right),$$

$$A(q, N) = q^{-s} \sum_{\substack{1 \leq a \leq q \\ (a, q)=1}} S(a, q)^s \exp \left(-2\pi i \frac{a}{q} N \right).$$

Then $\sum_{q=1}^{\infty} A(q, N)$ is absolutely convergent, and the sum is equal to $S(N)$. According to Hua (1959) we have

$$r_s(N) \sim S(N) \frac{\Gamma(1 + 1/k)^s}{\Gamma(s/k)} N^{s/k-1},$$

provided that $s \geq 2k^2(2 \log k + \log \log k + c)$. Next, if we denote by $V(N, \delta)$ the volume of the closed region satisfying $N \leq x_1^k + x_2^k + \dots + x_s^k \leq N + \delta$ in the s -dimensional Euclidean space, then $\lim_{\delta \rightarrow 0} V(N, \delta)/\delta = \chi_{\infty}(N)$ exists and is equal to $(\Gamma(1 + 1/k)^s / \Gamma(s/k)) N^{s/k-1}$. Hence we can show that the principal part of $r_s(N)$ is equal to the infinite product $\prod_p \chi_p(N)$, where p runs over all finite and infinite prime spots in \mathbf{Q} . This is a generalization of the singular series studied by Hardy.

With regard to $g(k)$, there are studies by L. E. Dickson (1936) and others. It is easy to see that $g(k) \geq 2k + (3/2)^k - 2$ and that $G(k) \geq k + 1$. It has been shown that $g(3) = 9$, $G(3) \leq 7$ (G. L. Watson, Linnik, 1947) and that $g(4) \leq 37$, $G(4) = 16$ (H. Davenport, 1939). More generally, Vinogradov (1959) proved that

$$G(k) \leq 2k \log k + 4k \log \log k + 2k \log \log \log k + ck.$$

To prove this, Vinogradov (1934) introduced the following integral, which is closely related to the prime number theorem:

$$I(P) = \int_0^1 \dots \int_0^1 \left| \sum_{x \leq P} \exp(2\pi i(\alpha_1 x + \alpha_2 x^2 + \dots + \alpha_k x^k)) \right|^{2s} d\alpha_1 \dots d\alpha_k.$$

Hua (1949) improved Vinogradov's result and proved that if $s \geq \frac{1}{4}k(k+1) + lk$, then

$$I(P) \leq (5s)^{5sl} (\log P)^{2l} P^{2s-k(k+1)/2+\delta},$$

where $\delta = \frac{1}{2}k(k+1)(1-1/k)^{l-1}$. Concerning $I(P)$, another notable approach was made by A. A. Karacuba and N. M. Korobov (1963). Further investigation proved that $I(P) = c_1 c_2 P^{2s-k(k+1)/2} + o(P^{2s-k(k+1)/2})$ if $s \geq ck^2 \log k$ (Vinogradov, Hua, 1959). The result is called the **Vinogradov mean value theorem**.

There are many variations and generalizations of this theorem. Vinogradov and Hua (1944) studied the problem of representing an arbitrary N as $N = p_1^k + p_2^k + \dots + p_s^k$ (with p_i prime). Hua (1937) and others also considered the problem of representing N as $N = f(x_1) + f(x_2) + \dots + f(x_s)$, where $f(x)$ is a given polynomial. Also, let

$$C(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n C_{ijk} x_i x_j x_k$$

be a homogeneous polynomial of degree 3 with integral coefficients. Davenport (1963) proved that if $n \geq 16$, then the equation $C(x_1, x_2, \dots, x_n) = 0$ has at least one nontrivial integral solution. There are further developments of the theory of representations of integers by forms in many variables by V. A. Tartakovskii, H. Davenport, B. J. Birch, and D. J. Lewis.

F. Additive Problems in an Algebraic Number Field

Siegel (1922) considered the generalization of Hardy's square sum problem to the case of an algebraic number field. He later (1945) studied the generalized Waring's problem in an algebraic number field K of finite degree: Let I be the principal order of K , and let J_k be the subring of I generated by the k th powers of integers in K . It is easily seen that the index $(I:J_k)$ is finite. Hence our concern regarding $s(k)$ must be restricted to integers contained in J_k . Another question is how to extend the Farey dissection to an algebraic number field. Siegel succeeded in solving these difficulties.

4 Ref.

Additive Number Theory

His ingenuity is seen in his way of dealing with the minor arcs, which provided a stimulus to the research of T. Mitsui (1960).

A generalization of Goldbach's problem to the case of an algebraic number field was obtained by Mitsui (1960) and O. Körner (1961).

As another extension of the Vinogradov three primes theorem, Mitsui (1971) proved the following theorem: Let K be an algebraic number field of degree n . Let C be the principal ideal class generated by a totally positive number in K , and P be the set of prime ideals of degree 1 contained in C . Let N be a positive integer and $I_s(N)$ be the number of representations of N as the sum of the norms of s prime ideals belonging to P ,

$$I_s(N) = \sum_{N = N_{p_1} + \dots + N_{p_s}} 1, \quad p_i \in P \quad (1 \leq i \leq s).$$

If N is sufficiently large and $s \geq 3$, we have the asymptotic formula

$$I_s(N) = A_s S(N) \frac{N^{s-1}}{(\log N)^s} + O\left(\frac{N^{s-1} \log \log N}{(\log N)^{s+1}} \right),$$

where A_s is a positive constant depending on s and K independent of N , and $S(N)$ denotes the singular series. If $s \equiv N \pmod{2D}$, where D is the discriminant of K , then $S(N) \geq c > 0$, where c is a constant. Later this problem was extended by Mitsui and T. Tatzuza (1981).

References

- [1] R. G. Ayoub, An introduction to the analytic theory of numbers, Amer. Math. Soc. Math. Surveys, 1963.
- [2] G. H. Hardy and E. M. Wright, An introduction to the theory of numbers, Clarendon Press, fourth edition, 1963.
- [3] H. H. Ostmann, Additive Zahlentheorie I, II, Erg. Math., Springer, 1956.
- [4] H. A. Rademacher, Lectures on analytic number theory, Lecture notes, Tata Inst., 1954–1955.
- [5] E. G. H. Landau, Über einige neuere Fortschritte der additiven Zahlentheorie, Cambridge Univ. Press, 1937.
- [6] L. K. Hua, Additive Primzahltheorie, Teubner, 1959.
- [7] L. K. Hua, Die Abschätzung von Exponentialsummen und ihre Anwendung in der Zahlentheorie, Enzykl. Math., Bd. I, 2, Heft 13, Teil I, 1959.
- [8] I. M. Vinogradov, Selected works (in Russian), Akad. Nauk SSSR, 1952.
- [9] I. M. Vinogradov, Upper boundary of $G(n)$ (in Russian), Izv. Akad. Nauk SSSR, 23 (1959), 637–642.
- [10] T. Mitsui, On the Goldbach problem in

an algebraic number field I, II, J. Math. Soc. Japan, 12 (1960), 290–324, 325–372.

[11] J.-R. Chen, On the representation of a large even integer as the sum of a prime and the product of at most two primes, Sci. Sinica, 16 (1973), 157–176.

[12] T. Mitsui and T. Tatzawa, Sur un problème en théorie additive des nombres, Monographies Scientifiques de la Maison Franco-Japonaise, fasc. 3, 1981.

5 (XVII.9) Additive Processes

A. General Remarks

The study of sums of \dagger independent random variables has been one of the main topics in modern probability theory (\rightarrow 250 Limit Theorems in Probability Theory). Combining the ideas of this study with the consideration of \dagger stochastic processes with continuous time parameter, we get the notion of additive processes.

B. Definitions and Fundamental Properties

A real-valued \dagger stochastic process $\{X_t\}_{0 \leq t < \infty}$, denoted by $X(t)$ ($0 \leq t < \infty$) for the rest of this article, where for simplicity we assume that $X(0)=0$, is called an **additive process** (or **process with independent increments**) if for any $t_0 < t_1 < \dots < t_n$, $X(t_i) - X(t_{i-1})$ ($i=1, 2, \dots, n$) are \dagger independent. An additive process is essentially the same as a \dagger spatially homogeneous \dagger Markov process (i.e., a Markov process on \mathbf{R}^1 that is invariant under translations). When, for any $h>0$ and $t>s$, $X(t+h) - X(s+h)$ and $X(t) - X(s)$ have the same law, i.e., the distribution law of $X(t) - X(s)$ depends only on $t-s$, we call the additive process $X(t)$ **temporally homogeneous**. This is essentially the same notion as a temporally and spatially homogeneous Markov process.

Let $X(t)$ be a given additive process. If $f(t)$ is a function of t only, then clearly $Y(t) = X(t) - f(t)$ is also an additive process, and we can choose $f(t)$ in such a way that for every $t>0$ and for every sequence $s_n \uparrow t$ ($s_n \downarrow t$), $Y(s_n)$ converges with probability one. Here, $\lim Y(s_n)$ is independent of a particular choice of s_n , and we denote it by $Y(t-) (=Y(t+))$. We call such $Y(t)$ a **central process** and also say that $Y(t)$ is obtained from $X(t)$ by **centering**. This $f(t)$ is given, for example, by the condition $E(\arctan(X(t) - f(t))) = 0$.

Let $Y(t)$ be a centered additive process. Then $Y(t-) = Y(t) = Y(t+)$ for all $t>0$, except on an at most countable t -set S , and $t \in S$ is called a **fixed point of discontinuity** of $Y(t)$. Then $Y_1(t) = \lim_{n \rightarrow \infty} U_n(t)$ exists with probability 1, where

$$U_n(t) = \sum_{\substack{0 \leq i \leq n \\ s_i < t}} (Y(s_i+) - Y(s_i-)) \\ + Y(t) - Y(t-) - C_t^n,$$

C_t^n being a constant determined by $E(\arctan U_n(t)) = 0$, and $S = \{s_j\}$ ($j=1, 2, \dots$). $Y_2(t) = Y(t) - Y_1(t)$ is a centered additive process without any fixed point of discontinuity. Furthermore, $Y_1(t)$ and $Y_2(t)$ are independent. Thus we have a decomposition of $Y(t)$: $Y(t) = Y_1(t) + Y_2(t)$, where $Y_1(t)$ and $Y_2(t)$ are mutually independent additive processes. The structure of $Y_1(t)$ is simple, and it is not worthwhile to study its behavior in more detail. On the other hand, since $Y_2(t)$ is a centered additive process without any fixed point of discontinuity, it is an additive process that is \dagger continuous in probability. Let $\tilde{Y}_2(t)$ be a \dagger separable modification of $Y_2(t)$. Then the discontinuities of almost all sample functions of $\tilde{Y}_2(t)$ are of the first kind. If we set $Y_2^*(t) = \tilde{Y}_2(t+)$, $Y_2^*(t)$ is a \dagger modification of $Y_2(t)$, and almost all sample functions of $Y_2^*(t)$ are right continuous and have a left-hand limit at every t . In the study of the process $Y_2(t)$, it is always convenient to take such a modification. Thus we give the following general definition: An additive process is called a **Lévy process** if it is continuous in probability and almost all sample functions are right continuous and have a left-hand limit at every $t \in [0, \infty)$ [3, 9].

The notions of additive processes and Lévy processes can also be considered for \mathbf{R}^N -valued processes.

C. Additive Processes and Infinitely Divisible Distributions

Let $X(t)$ be a Lévy process and Φ_{st} ($s < t$) be the \dagger distribution of $X(t) - X(s)$. Then Φ_{st} is an \dagger infinitely divisible distribution (\rightarrow 341 Probability Measures G). Conversely, for a given infinitely divisible distribution Φ we can construct an essentially unique temporally homogeneous Lévy process $X(t)$ such that Φ coincides with the distribution of $X(1)$. If $X(t)$ is temporally homogeneous, the \dagger characteristic function $\varphi_{st}(z) = E(e^{iz(X(t) - X(s))})$ of the distribution Φ_{st} is given in the form $\varphi_{st} = \exp((t-s)\psi(z))$; hence the law of the process $X(t)$ is completely determined by the function $\psi(z)$. By the \dagger Lévy-Khinchin canonical form, $\psi(z)$ is

written in the form

$$\psi(z) = imz - \frac{v}{2}z^2 + \int_{-\infty}^{\infty} \left(e^{izu} - 1 - \frac{izu}{1+u^2} \right) n(du), \quad (1)$$

where $m, v \in \mathbf{R}$, $v \geq 0$, and $n(du)$ is a nonnegative measure on $\mathbf{R} - \{0\}$ such that $\int_{-\infty}^{\infty} \frac{u^2}{1+u^2} n(du) < \infty$. These m, v , and $n(du)$ are uniquely determined by $\psi(z)$.

D. Basic Additive Processes

Wiener Process. When almost all sample functions of a Lévy process $X(t)$ are continuous, the distribution of $X(t) - X(s)$ is a \dagger normal distribution. If, further, $X(t)$ is temporally homogeneous, $\psi(z)$ has the form $\psi(z) = imz - \frac{1}{2}vz^2$. In particular, if $m = 0$ and $v = 1$, then $X(t)$ is called a **Wiener process** or **Brownian motion**. This stochastic process was introduced by N. Wiener (1923) as a mathematical model for the random movement of colloidal particles first observed by a British botanist, R. Brown. This is one of the most fundamental and important stochastic processes in modern probability theory (\rightarrow 45 Brownian Motion).

Poisson Processes. When almost all sample functions of a Lévy process are increasing step functions with only jumps of size 1, the distribution of $X(t) - X(s)$ is a \dagger Poisson distribution. If, further, $X(t)$ is temporally homogeneous, $\psi(z)$ in (1) has the form $\psi(z) = \lambda(e^{iz} - 1)$ ($\lambda > 0$), and $X(t)$ is called a **Poisson process**. Let $X(t)$ be a Poisson process, and let $T_0, T_0 + T_1, T_0 + T_1 + T_2, \dots$ be successive jumping times of a sample function $X(t)$. Then T_0, T_1, T_2, \dots is a sequence of mutually independent random variables with the common exponential distribution $P(T \in dt) = \lambda e^{-\lambda t} dt$. Conversely, given such a sequence $\{T_n\}$, if we define $X(t) = \inf\{n | T_0 + T_1 + \dots + T_n > t\}$, then $X(t)$ is a Poisson process. Thus, for example, the number of telephone calls at a switchboard is a Poisson process when the intervals between successive calls can be regarded as independent and having a common exponential distribution.

E. The Structure of the General Lévy Process [8, 9]

In this section we restrict ourselves for simplicity to temporally homogeneous Lévy processes. As we noted, the probability law of the process $X(t)$ is determined by the function $\psi(z)$. The Lévy-Khinchin formula (1) in a

5 E

Additive Processes

certain sense shows that ψ is a combination of a Wiener process and Poisson processes. This fact can be seen more clearly from the **Lévy-Itô theorem**, which states that the sample function of $X(t)$ itself can be expressed as a composite of those of a Wiener process and Poisson processes. The Lévy-Itô theorem actually implies formula (1) and, moreover, clarifies its probabilistic meaning.

The Lévy-Itô theorem can be summarized as follows: Let U be a Borel subset of \mathbf{R} which has a positive distance from the origin, and let $N(t, U)$ be the number of s such that $X(s) - X(s-) \in U$, $s \leq t$. Then $N(t, U)$ is a Poisson process. The expectation $E(N(t, U))$ can be written in the form $tn(U)$, where $n(U)$ defines a nonnegative Borel measure on $\mathbf{R} - \{0\}$; furthermore, it satisfies

$$\int_{-\infty}^{\infty} \frac{u^2}{1+u^2} n(du) < \infty.$$

Next, we set

$$\begin{aligned} S_\varepsilon(t) &= \int_{|u| > \varepsilon} u N(t, du) \\ &= \sum_{\substack{s \leq t \\ |X(s) - X(s-)| \geq \varepsilon}} (X(s) - X(s-)). \end{aligned}$$

Generally, $S_\varepsilon(t)$ diverges as $\varepsilon \downarrow 0$. However, with probability one, a centered process

$$\bar{S}_\varepsilon(t) = S_\varepsilon(t) - t \int_{|u| > \varepsilon} \frac{u}{1+u^2} n(du)$$

converges uniformly in t on every finite interval as $\varepsilon \downarrow 0$. Furthermore, $X(t) - \lim_{\varepsilon \downarrow 0} \bar{S}_\varepsilon(t)$ is continuous with probability one. However, a Lévy process $X(t)$ of which almost all sample functions are continuous has the form $mt + \sqrt{v} B(t)$, where $m, v \geq 0$ are constants and $B(t)$ is a Wiener process. Hence we have

$$\begin{aligned} X(t) &= mt + \sqrt{v} B(t) \\ &+ \lim_{\varepsilon \downarrow 0} \int_{|u| > \varepsilon} \left(u N(t, du) - \frac{u}{1+u^2} tn(du) \right). \end{aligned} \quad (2)$$

Furthermore, we can show that if

$$U_1, U_2, \dots, U_n$$

are disjoint, then

$$B(t), N(t, U_1), N(t, U_2), \dots, N(t, U_n)$$

are mutually independent Lévy processes. In particular, in (2) the terms are mutually independent. The m, v , and $n(du)$ in (2) correspond, of course, to those in (1). Conversely, given m, v , and $n(du)$, we can construct $B(t)$ and $N(t, U)$ with the properties above, and then (2) defines a Lévy process which corresponds to $\psi(z)$ given by (1). The measure $n(du)$ is called the **Lévy measure** of $X(t)$.

F. Examples of Lévy Processes

Compound Poisson Processes. A temporally homogeneous Lévy process is called a **compound Poisson process** if almost all sample functions are step functions, namely, if $\psi(z)$ in (1) is given by

$$\psi(z) = \int_{-\infty}^{\infty} (e^{izu} - 1)n(du),$$

$$\lambda = \int_{-\infty}^{\infty} n(du) < \infty.$$

If we set $\Phi(du) = (1/\lambda)n(du)$, then $\Phi(du)$ is a probability distribution on \mathbf{R} . Φ is the distribution of the size of jumps when they occur. A compound Poisson process is constructed in the following way: Let $T_0, T_1, T_2, \dots; U_1, U_2, \dots$ be mutually independent random variables such that

$$P(T_n \in dt) = \lambda e^{-\lambda t} dt, \quad t > 0,$$

$$P(U_n \in du) = \Phi(du),$$

and let

$$N(t) = U_1 + U_2 + \dots + U_{N(t)},$$

where

$$n = \inf\{n | T_0 + T_1 + \dots + T_n > t\}.$$

Then $X(t)$ is a compound Poisson process. Thus the number of jumps of $X(t)$ follows a Poisson process, and the size of each jump obeys the distribution Φ .

Stable Processes. A temporally homogeneous Lévy process $X(t)$ is called a **stable process** if for every $a > 0$ we can find $b > 0$ and c real such that the processes $X_1(t) = X(at)$ and $X_2(t) = bX(t) + ct$ are equivalent in law. It is called a **strictly stable process** if in the above c can be chosen to be 0. $X(t)$ is a stable process (resp. a strictly stable process) if and only if the corresponding infinitely divisible distribution is a † quasistable distribution (resp. † stable distribution). The † exponent α ($0 < \alpha \leq 2$) of the quasistable distribution is called the **exponent** (or **index**) of the stable process. $\psi(z)$ in (1) corresponding to a stable process is given as follows:

$$\psi(z) = imz + C_+ \int_0^{\infty} (e^{izu} - 1) \frac{du}{u^{1+\alpha}}$$

$$+ C_- \int_{-\infty}^0 (e^{izu} - 1) \frac{du}{|u|^{1+\alpha}}, \quad 0 < \alpha < 1,$$

$$\psi(z) = imz + C_+ \int_0^{\infty} \left(e^{izu} - 1 - \frac{izu}{1+u^2} \right) \frac{du}{u^2}$$

$$+ C_- \int_{-\infty}^0 \left(e^{izu} - 1 - \frac{izu}{1+u^2} \right) \frac{du}{|u|^2}, \quad \alpha = 1,$$

$$\psi(z) = imz + C_+ \int_0^{\infty} (e^{izu} - 1 - izu) \frac{du}{u^{1+\alpha}}$$

$$+ C_- \int_{-\infty}^0 (e^{izu} - 1 - izu) \frac{du}{|u|^{1+\alpha}}, \quad 1 < \alpha < 2,$$

$$\psi(z) = imz - \frac{v}{2} z^2, \quad \alpha = 2.$$

Here m, C_+, C_- , and v are real constants such that $C_+ \geq 0, C_- \geq 0$ with $C_+ + C_- > 0$ and $v > 0$. $\psi(z)$ corresponds to a strictly stable process if and only if $m=0$ when $\alpha \neq 1$ and $C_+ = C_-$ when $\alpha = 1$. The above $\psi(z)$ is also expressed as follows:

(i) if $0 < \alpha \leq 2, \alpha \neq 1$, then

$$\psi(z) = imz - c_0 |z|^{\alpha} \left(1 - i\beta \tan \frac{\pi\alpha}{2} \frac{z}{|z|} \right),$$

where

$$c_0 = \frac{v}{2} \quad \text{if } \alpha = 2$$

and

$$c_0 = -(C_+ + C_-) \Gamma(-\alpha) \cos \frac{\pi\alpha}{2},$$

$$\beta = \frac{C_+ - C_-}{C_+ + C_-} \quad \text{if } \alpha \neq 2;$$

(ii) if $\alpha = 1$, then

$$\psi(z) = i\gamma z - c_0 |z| \left(1 + i\frac{2}{\pi} \beta \frac{z}{|z|} \log |z| \right),$$

where

$$\gamma = m + a(C_+ + C_-)$$

$$\left(a = \int_0^{\infty} \left(\sin u - \frac{u}{1+u^2} \right) \frac{du}{u^2} \right),$$

$$c_0 = \frac{\pi}{2} (C_+ - C_-), \quad \beta = \frac{C_+ - C_-}{C_+ + C_-}.$$

When $\alpha = 2$, it is thus essentially a Wiener process: $X(t) = mt + \sqrt{v} B(t)$, where $B(t)$ is a Wiener process. When $\alpha = 1$, it is called a **Cauchy process**, a **symmetric Cauchy process** if $m=0$ and $C_+ = C_-$ or equivalently $\gamma = \beta = 0$, or an **asymmetric Cauchy process** if $\beta \neq 0$. Generally it is called a **symmetric stable process** if $m=0$ and $C_+ = C_-$ or $\alpha = 2$ and $m=0$. In particular, for the symmetric Cauchy process corresponding to $\psi(z) = -|z|$, we have

$$P(X(t) < x) = \frac{1}{\pi} \int_{-\infty}^x \frac{t}{t^2 + y^2} dy.$$

Next, when $0 < \alpha < 1$ and $m = C_- = 0$, almost all sample functions of $X(t)$ are purely discontinuous increasing functions (i.e., sums of positive jumps). In this case, $X(t)$ is called a **one-sided stable process of the exponent α** (or **subordinator of the exponent α**). Now, if $X(t)$ is

a symmetric stable process of the exponent β ($0 < \beta \leq 2$) and $Y(t)$ is a subordinator of the exponent α which is independent of $X(t)$, then $Z(t) = X(Y(t))$ is a symmetric stable process of the exponent $\alpha\beta$. This operation is called a **subordination** and is closely related to the theory of fractional powers of infinitesimal generators of semigroups (\rightarrow 261 Markov Processes) [2].

A stable process $X(t)$ is defined in a similar way when $X(t)$ takes values in an N -dimensional space \mathbf{R}^N . In particular, if $X(t)$ is a symmetric stable process of the exponent α ($0 < \alpha \leq 2$) given by

$$E(e^{i(z, X(t))}) = e^{-t|z|^\alpha}, \quad z \in \mathbf{R}^N,$$

then for every bounded measurable function $f(x)$ with compact support, we have

$$\begin{aligned} E\left(\int_0^\infty f(x + X(t)) dt\right) \\ = \frac{\Gamma((N-\alpha)/2)}{2^\alpha \pi^{N/2} \Gamma(\alpha/2)} \int_{\mathbf{R}^N} |x-y|^{\alpha-N} f(y) dy. \end{aligned}$$

The right-hand side is the \dagger Riesz potential of order α (\rightarrow 338 Potential Theory). This fact is a generalization of a well-known relation between Brownian motion and Newtonian potential (\rightarrow 45 Brownian Motion), and through this relation we can study several properties of sample functions and also compute various quantities related to stable processes [1].

G. Sample Path Properties of Lévy Processes

Let $X(t)$ be a temporally homogeneous Lévy process. For a Borel set B in \mathbf{R} , the hitting time σ_B is defined by

$$\sigma_B = \inf\{t > 0 \mid X(t) \in B\}.$$

Recurrence. $X(t)$ is called **recurrent** if $\sigma_B < \infty$ a.s. for every nonempty open set B . Otherwise it is called **transient**. $X(t)$ is recurrent if and

only if $\int_{-1}^1 \operatorname{Re} \frac{1}{-\psi(z)} dz = \infty$ (Port and Stone, Ornstein). If, in particular, $E(X(1))$ exists, it is recurrent if and only if $E(X(1)) = 0$. When $X(t)$ is a stable process, it is recurrent if and only if $\alpha > 1$, or $\alpha = 1$ and $\beta = 0$.

Hitting Probabilities for Single Points. If $B = \{a\}$, σ_B is denoted by σ_a . $0 \in \mathbf{R}$ is said to be **regular** for $X(t)$ if $\sigma_0 = 0$ a.s. Set $C = \{x \in \mathbf{R} \mid P(\sigma_x < \infty) > 0\}$. The following result is due to Kesten [10]:

- (i) If $v \neq 0$, then $C = \mathbf{R}$ and 0 is regular;
- (ii) If $v = 0$ and $\int_{-1}^1 |u| n(du) = \infty$, then either $C = \mathbf{R}$ and 0 is regular, or $C = \emptyset$ and 0 is not

regular according to whether $\int_{\mathbf{R}} \operatorname{Re} \frac{1}{\lambda - \psi(z)} dz < \infty$ for all $\lambda > 0$ or $= \infty$ for some $\lambda > 0$;

(iii) If $v = 0$ and $\int_{-1}^1 |u| n(du) < \infty$, then the following four cases are possible, where $a = m - \int_{\mathbf{R}} u(1+u^2)^{-1} n(du)$ and $S(n)$ is the support of the Lévy measure n :

- (a) $C = \emptyset$ when $a = 0$,
- (b) $C = (0, \infty)$ when $a > 0$ and $S(n) \subset (0, \infty)$,
- (c) $C = (-\infty, 0)$ when $a < 0$ and $S(n) \subset (-\infty, 0)$, and
- (d) $C = \mathbf{R}$ in the remaining case.

For further properties of sample functions \rightarrow [4].

H. Generalization of Additive Processes

A temporally homogeneous Lévy process is, as we have seen, essentially a temporally homogeneous Markov process on \mathbf{R} which is homogeneous in space (i.e., invariant under translations of the space). Thus, on a homogeneous space when homogeneity in space makes sense, we can generalize the notion of additive processes. Let M be a \dagger homogeneous space with transformation group G . A temporally homogeneous Markov process $X(t)$ is called an **invariant Markov process**, (or **homogeneous Markov process**) if its system of \dagger transition probabilities $\{P(t, x, E)\}$ satisfies $P(t, x, E) = P(t, gx, gE)$ for all $g \in G$. Thus an additive process is exactly an invariant Markov process on \mathbf{R}^N when G is the group of translations. G. A. Hunt determined all invariant Markov processes when M is a \dagger Lie group or a factor space of a Lie group [6].

Let G be a Lie group and $\Lambda = \Lambda(G)$ be the (left-invariant) \dagger Lie algebra of G . Let $G_c = G \cup \{\omega\}$ be a \dagger one-point compactification of G , and C be the set of all continuous functions on G_c . We can define $Y(f) (Y \in \Lambda, f \in C)$ as usual by

$$Y(f) = \lim_{t \downarrow 0} \frac{R_{\eta(t)} f - f}{t},$$

$$\eta(t) = \exp t Y, \quad R_\sigma f(\tau) = f(\tau\sigma),$$

when the limit exists uniformly. Let $C_2 = \{f \in C \mid Y(Xf) \text{ exists for every } X, Y \in \Lambda\}$. Let X_1, X_2, \dots, X_d be a basis of $\Lambda(G)$, and let x_1, x_2, \dots, x_d be functions in C_2 such that $x_i(e) = 0$ and $X_i(x_j)(e) = \delta_{ij}$ ($i, j = 1, 2, \dots, d$; e is the unit element of G). Take a neighborhood of e , and define a function $\varphi(g) = \sum_{i=1}^d x_i^2(g)$ for g contained in the neighborhood, and extend this function to G_c in such a way that $\varphi \in C_2$ and $\varphi \geq k > 0$ ($k = \text{constant}$) outside of the neighborhood of e . Then $g \in G$ defines a transformation of G_c by $\tau_g \sigma = g\sigma$, $\tau_g \omega = \omega$, and in this way G_c is supplied with the structure of a \dagger homogeneous space with the transformation

group G . Now let $X(t)$ be an invariant Markov process on G_c which is \dagger continuous in probability. Then the \dagger semigroup T_t (which is a \dagger strongly continuous semigroup on C) of the process $X(t)$ is characterized as follows: The domain of the infinitesimal generator A of T_t contains C_2 , and for $f \in C_2$

$$\begin{aligned} Af(\tau) = & \sum_{i=1}^d a_i(X_i f)(\tau) + \sum_{i,j=1}^d a_{ij} X_i(X_j f)(\tau) \\ & + \int_{G_c - \{e\}} \left\{ f(\tau\sigma) - f(\tau) \right. \\ & \left. - \sum_{i=1}^d X_i f(\tau) \cdot x_i(\sigma) \right\} n(d\sigma), \end{aligned}$$

where a_i, a_{ij} are real numbers ($i, j = 1, 2, \dots, d$) such that (a_{ij}) is a symmetric nonnegative definite matrix, and $n(d\sigma)$ is a nonnegative measure on $G_c - \{e\}$ such that $\int_{G_c - \{e\}} \varphi(\sigma) n(d\sigma) < \infty$. Conversely, given such a_i, a_{ij} and $n(d\sigma)$, there exists one and only one invariant Markov process on G_c whose infinitesimal generator is given as above.

A similar result is obtained when M is a factor space of a Lie group by its compact subgroup. Furthermore, for more concrete homogeneous spaces such as spheres or Lobachevskii spaces (more generally, \dagger symmetric Riemannian spaces) the canonical form of the invariant Markov processes and infinitely divisible laws is obtained by making use of harmonic analysis [5].

References

- [1] R. M. Blumenthal, R. K. Gettoor, and D. B. Ray, On the distribution of the first hits for the symmetric stable processes, *Trans. Amer. Math. Soc.*, 99 (1961), 540–554.
- [2] S. Bochner, *Harmonic analysis and the theory of probability*, Univ. of California Press, 1955.
- [3] J. L. Doob, *Stochastic processes*, Wiley, 1953.
- [4] B. Fristedt, Sample functions of stochastic processes with stationary, independent increments, *Advances in Probability and Related Topics III*, P. Ney and S. Port (eds.), Dekker, 1974, 241–396.
- [5] R. Gangolli, Isotropic infinitely divisible measures on symmetric spaces, *Acta Math.*, 111 (1964), 213–246.
- [6] G. A. Hunt, Semi-groups of measures on Lie groups, *Trans. Amer. Math. Soc.*, 81 (1956), 264–293.
- [7] K. Itô, On stochastic processes I, *Japan. J. Math.*, 18 (1942), 261–301.
- [8] K. Itô, On stochastic processes, *Lecture notes*, Tata Inst., 1960.

- [9] K. Itô, *Stochastic processes*, Lecture Notes Series, no. 16, Aarhus Univ., 1969.
- [10] H. Kesten, Hitting probabilities of single points for processes with stationary independent increments, *Mem. Amer. Math. Soc.*, 93 (1969).
- [11] P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, 1937.
- [12] P. Lévy, *Processus stochastiques et mouvement brownien*, Gauthier-Villars, 1948.

6 (III.20) Adeles and Ideles

A. Introduction

The concept of idele was first introduced by C. Chevalley (*J. Math. Pures Appl.*, (9) 15 (1936); *Ann. Math.*, (2) 41 (1940)), for \dagger algebraic number fields. Later on, this concept and the allied concept of adele were defined for \dagger simple algebras and also for \dagger algebraic groups over algebraic number fields, and the two concepts became important in the arithmetical theory of these objects. We shall first explain the general concept of restricted direct product, by means of which adeles and ideles will be defined.

B. Restricted Direct Product

Let I be an index set. Suppose we are given, for each $p \in I$, a \dagger locally compact group G_p , and for each p except for a given finite set, say p_1, p_2, \dots, p_r , a compact open subgroup U_p of G_p . Let G be the subgroup of the direct product $\prod_{p \in I} G_p$ consisting of elements (g_p) whose G_p -components g_p lie in U_p , except for a finite number of p . Put $U = \prod_{i=1}^r G_{p_i} \times \prod_{p \neq p_i} U_p$. Then U is a locally compact group with respect to the \dagger product topology. The group G can be supplied naturally with a topology with respect to which G is a locally compact group and the quotient space G/U is discrete. The group G together with this topology is called the **restricted direct product** of $\{G_p\}$ with respect to $\{U_p\}$.

C. Adeles and Ideles

Let k be an \dagger algebraic number field of finite degree and I be the totality of finite and infinite \dagger prime divisors of k . For each $p \in I$ we denote by k_p and k_p^* the \dagger completion of k with respect to p and the multiplicative group of nonzero elements of k_p , respectively. Furthermore, for each finite prime divisor p , we denote by \mathfrak{o}_p and u_p the ring of $\dagger p$ -adic integers

of k and the multiplicative group of \dagger units of \mathfrak{o}_p , respectively.

(1) Since \mathfrak{o}_p is a compact open subgroup of k_p as an additive group, we can construct the restricted direct product \mathbf{A}_k of $\{k_p\}$ with respect to $\{\mathfrak{o}_p\}$. Then \mathbf{A}_k is a locally compact ring with respect to the componentwise ring operations. We call \mathbf{A}_k the **adele ring** (or **ring of valuation vectors**) of k , and an element of \mathbf{A}_k an **adele** (or **valuation vector**) of k . The element of the direct product $\prod_p k_p$ whose p -component is a fixed element of k for all p is an adele. We call such an adele a **principal adele**. Since \mathfrak{o}_p is a compact open subgroup of k_p^\times for each finite prime p , we can construct the restricted direct product \mathbf{J}_k of $\{k_p^\times\}$ with respect to $\{\mathfrak{o}_p\}$. We call \mathbf{J}_k the **idele group** of k and an element of \mathbf{J}_k an **idele** of k . The element of the direct product $\prod_p k_p^\times$ whose p -component is a fixed element of k for all p is an idele. We call such an idele a **principal idele**. Each element b of \mathbf{J}_k induces an automorphism f_b of the additive group \mathbf{A}_k defined by $f_b(a) = b \cdot a$ ($a \in \mathbf{A}_k$). Thus \mathbf{J}_k can be regarded as a subgroup of the automorphism group $\text{Aut}(\mathbf{A}_k)$ of the additive group \mathbf{A}_k . The topology of \mathbf{J}_k coincides with the relative topology of \mathbf{J}_k as a subgroup of $\text{Aut}(\mathbf{A}_k)$. We note, however, that the topology of \mathbf{J}_k is different from the relative topology of \mathbf{J}_k as a subspace of \mathbf{A}_k , and the former is stronger than the latter. Finally, for a \dagger function field in one variable over a \dagger finite field, the adele ring and the idele group can be defined similarly.

(2) Let \mathfrak{R} be a \dagger normal simple algebra over k and \mathfrak{O} be a \dagger maximal order of \mathfrak{R} . For each $p \in I$ put $\mathfrak{R}_p = \mathfrak{R} \otimes_k k_p$, and for each finite prime divisor p put $\mathfrak{O}_p = \mathfrak{o}_p \cdot \mathfrak{O}$. Then \mathfrak{O}_p is a compact open additive subgroup of \mathfrak{R}_p . By the adele ring $\mathbf{A}_{\mathfrak{R}}$ of \mathfrak{R} we mean the restricted direct product of $\{\mathfrak{R}_p\}$ with respect to $\{\mathfrak{O}_p\}$. Let \mathfrak{R}_p^\times and \mathfrak{U}_p be the multiplicative group of nonzero divisors of \mathfrak{R}_p and the multiplicative group of the units of \mathfrak{O}_p , respectively. (\mathfrak{U}_p can be defined only if p is a finite prime divisor.) By the idele group $\mathbf{J}_{\mathfrak{R}}$ of \mathfrak{R} we mean the restricted direct product of $\{\mathfrak{R}_p^\times\}$ with respect to $\{\mathfrak{U}_p\}$. The notion of principal adele (or idele) of \mathfrak{R} can be defined similarly, as in (1). The structures, as topological groups, of $\mathbf{A}_{\mathfrak{R}}$ and $\mathbf{J}_{\mathfrak{R}}$ do not depend on the choice of a maximal order \mathfrak{O} . The adele ring \mathbf{A}_k and the idele group \mathbf{J}_k described in (1) are special cases of $\mathbf{A}_{\mathfrak{R}}$ and $\mathbf{J}_{\mathfrak{R}}$, respectively.

(3) Let G be a linear \dagger algebraic group defined over k , and let G_p be the set of k_p - \dagger rational points of the group for each $p \in I$. For each finite prime divisor p , let U_p be the set of elements α of G_p such that the coordinates of both α and α^{-1} are p -adic integers. We can then construct the restricted direct product of

$\{G_p\}$ with respect to $\{U_p\}$, which is called the **idele group** (or **adele group**) of G .

In the following section we focus on describing the fundamental properties of adeles and ideles of an algebraic number field k . We shall start, however, by observing more generally those adeles and ideles of a normal simple algebra \mathfrak{R} over k . (For the properties of the adele group of algebraic groups \rightarrow [7], 13 Algebraic Groups.)

D. The Structures of the Adele Ring and Idele Group

Let \mathfrak{R} be a normal simple algebra over an algebraic number field of finite degree k . We identify the totality of principal adeles of \mathfrak{R} (principal ideles of \mathfrak{R}) with $\mathfrak{R}(\mathfrak{R}^\times)$, and denote it by the same letter $\mathfrak{R}(\mathfrak{R}^\times)$. Then $\mathfrak{R}(\mathfrak{R}^\times)$ is a discrete subgroup of $\mathbf{A}_{\mathfrak{R}}(\mathbf{J}_{\mathfrak{R}})$. The quotient group $\mathbf{A}_{\mathfrak{R}}/\mathfrak{R}$ is compact. Denoting by $|\alpha_p|_p$ ($\alpha_p \in k_p$) and $N_p(\alpha_p)$ ($\alpha_p \in \mathfrak{R}_p$) the \dagger normalized valuation of k_p and the \dagger reduced norm from \mathfrak{R}_p to k_p , respectively, we define, for $a \in \mathbf{J}_{\mathfrak{R}}$, a positive number: $V(a) = \prod_{p \in I} |N_p(\alpha_p)|_p$, where $a = (\alpha_p)$. We call $V(a)$ the **volume** of a . If a is a principal idele, we have $V(a) = 1$ by the \dagger product formula on valuations. Denote by $\mathbf{J}_{\mathfrak{R}}^0$ the set of ideles a with $V(a) = 1$ and put $\mathbf{C}_{\mathfrak{R}}^0 = \mathbf{J}_{\mathfrak{R}}^0/\mathfrak{R}^\times$. Then $\mathbf{C}_{\mathfrak{R}}^0$ has finite volume with respect to the \dagger Haar measure of $\mathbf{J}_{\mathfrak{R}}$. Furthermore, $\mathbf{C}_{\mathfrak{R}}^0$ is compact if and only if \mathfrak{R} is a \dagger division algebra. In particular, \mathbf{C}_k^0 is compact. Let \mathbf{Q} be the field of rational numbers. For each rational prime p , we define a character λ_p of the completion \mathbf{Q}_p of \mathbf{Q} with respect to the p -adic topology (by a character of \mathbf{Q}_p , we mean a continuous homomorphism from \mathbf{Q}_p to the 1-dimensional torus \mathbf{R}/\mathbf{Z}): If $p = p_\infty$ is the infinite prime of \mathbf{Q} , then we put $\lambda_{p_\infty}(x) \equiv -x \pmod{\mathbf{Z}}$ ($x \in \mathbf{Q}_p$). If p is finite, then we let λ_p be the composite of the following three canonical homomorphisms, namely, the one from \mathbf{Q}_p to $\mathbf{Q}_p/\mathbf{Z}_p$, the one from $\mathbf{Q}_p/\mathbf{Z}_p$ to \mathbf{Q}/\mathbf{Z} , and the one from \mathbf{Q}/\mathbf{Z} to \mathbf{R}/\mathbf{Z} . We define a character λ_p of \mathfrak{R}_p as follows: $\lambda_p = \lambda_p \circ \text{tr}(\mathfrak{R}_p/\mathbf{Q}_p)$, where p is the rational prime divisible by p and $\text{tr}(\mathfrak{R}_p/\mathbf{Q}_p)$ denotes the \dagger reduced trace from \mathfrak{R}_p to \mathbf{Q}_p . For $x, y \in \mathfrak{R}_p$, put $(x, y)_p = \exp(2\pi i \lambda_p(xy))$. Then the additive group \mathfrak{R}_p is self-dual relative to $(x, y)_p$. Furthermore, if we put $\langle a, b \rangle = \prod_p (a_p, b_p)_p$ for $a = (a_p)$ and $b = (b_p) \in \mathbf{A}_{\mathfrak{R}}$ then $\mathbf{A}_{\mathfrak{R}}$ is self-dual relative to $\langle a, b \rangle$. The \dagger annihilator of the group of principal adeles with respect to $\langle a, b \rangle$ is \mathfrak{R} . Hence it follows from \dagger Pontryagin's duality theorem that $\mathbf{A}_{\mathfrak{R}}/\mathfrak{R}$ is compact. Henceforth let $\mathfrak{R} = k$. We call the quotient group $\mathbf{C}_k = \mathbf{J}_k/k^\times$ (an element of \mathbf{C}_k) the **idele class group** of k (an **idele class**). If a character χ of \mathbf{J}_k satisfies the condition $\chi(\alpha) = 1$ for all $\alpha \in k$ (i.e.,

if χ is a character of C_k , we call such a character a **Größencharakter** (or **Hecke character**). Größencharakteren were introduced by E. Hecke as characters of a certain type of the ideal group of k (*Math. Z.*, 1 (1918), 5 (1920)), but they are essentially the same as the ones defined above [1]. Let D_k be the connected component of the identity element of C_k . Then C_k/D_k is totally disconnected and compact. Hence a Größencharakter χ is of finite order if and only if $\chi(D_k) = 1$. We can prove by class field theory that C_k/D_k is canonically isomorphic to the Galois group over k of the maximal Abelian extension of k (\rightarrow 59 Class Field Theory). For the structure of D_k , the following fact is known: Let r_1 and r_2 be the number of real infinite prime divisors and imaginary infinite prime divisors of k , respectively. Then the dual group of D_k is isomorphic to $\mathbf{R} \times \mathbf{Q}^{r_1+r_2-1} \times \mathbf{Z}^{r_2}$, where \mathbf{R} is the additive group of real numbers with the usual topology and $\mathbf{Q}(\mathbf{Z})$ is the additive group of rational numbers (rational integers) with the discrete topology. Let F be a function field in one variable over a finite field \mathbf{F}_0 . The properties of the adèle ring and idele group of F are similar to the properties of A_k and J_k , while the group C_F has a simpler structure than C_k . To explain the structure of C_F , let \tilde{F} be the maximal Abelian extension of F , \tilde{G} be the Galois group of \tilde{F}/F , and G_F be the subgroup of \tilde{G} consisting of the elements σ such that $\sigma(\alpha) = \alpha^{q^n}$ for all $\alpha \in \bar{\mathbf{F}}_0$ (= the algebraic closure of \mathbf{F}_0), where q is the number of elements of the finite field \mathbf{F}_0 and n is a given rational integer. Also, let G_F^0 be the subgroup of G_F consisting of the elements inducing the identity map on $\bar{\mathbf{F}}_0$. G_F^0 is a compact group with respect to the Krull topology. G_F can be naturally supplied with a topology such that the group G_F is a locally compact group and the quotient group G_F/G_F^0 is discrete. Then class field theory implies that G_F is isomorphic to C_F as a topological group.

The following characterization of the adèle ring of a number field or function field in one variable over a finite field is the work of K. Iwasawa (*Ann. Math.*, (2) 57 (1953)). Let A be a semisimple commutative and locally compact topological ring with unity 1. Assume that A is neither discrete nor compact, and moreover that A contains a discrete subfield $k \ni 1$ and A/k is compact. Then k is an algebraic number field or a function field in one variable over a finite field, and A is isomorphic to the adèle ring of k as a topological ring.

E. Ideles and Cohomology

Let K be a Galois extension of finite degree of an algebraic number field k , and \mathfrak{G} be the

Galois group of the extension K/k . \mathfrak{G} operates naturally on the idele group J_K and the idele class group C_K of K . The structures of the cohomology groups of \mathfrak{G} with the coefficient groups J_K and C_K were investigated by G. Hochschild, T. Nakayama, E. Artin, J. Tate, and others. In particular, we have $H^1(\mathfrak{G}, C_K) = \{0\}$ and $H^2(\mathfrak{G}, C_K) \cong \mathbf{Z}/n\mathbf{Z}$ (cyclic group of order n), where $n = [K:k]$. These facts play an important role in one of the proofs of class field theory (\rightarrow [3]; 59 Class Field Theory). Furthermore, A. Weil introduced the so-called **Weil group**, which is a group extension of a certain type of C_K by \mathfrak{G} . He defined the most general L -functions, which include both Artin L -functions and Hecke L -functions with Größencharakteren (\rightarrow [2]; 450 Zeta Functions).

F. Fourier Analysis on the Adele Group

Dedekind zeta functions and Hecke L -functions are meromorphic on the whole complex plane and satisfy functional equations of certain types. This can be proved by methods of Fourier analysis on the adèle group A_k (Artin, Iwasawa, Tate [1, 8, 9]). For a continuous complex-valued function $\varphi(a)$ on A_k satisfying suitable conditions, we define the Fourier transform of $\varphi(a)$ as follows:

$$\hat{\varphi}(a) = \int_{A_k} \varphi(b) \langle a, b \rangle db,$$

where db denotes the Haar measure on A_k . By normalizing db suitably and applying Poisson's summation formula, we get, for each idele a of k ,

$$\sum_{\alpha \in k} \varphi(a\alpha) = V(a)^{-1} \sum_{\alpha \in k} \hat{\varphi}(a^{-1}\alpha).$$

This is called the **Θ -formula**. Consider the following integral on J_k :

$$\xi(s) = \int_{J_k} V(a)^s \chi(a) \varphi(a) d^*a,$$

where d^*a denotes the Haar measure on J_k , s is a complex number, and χ is a Größencharakter of k , namely, a character of C_k . This integral converges if $s > 1$, and by using the Θ -formula one can show that $\xi(s)$ is meromorphic on the whole complex plane and satisfies a functional equation of a certain type. When the function φ is of special type, then the above integral can be explicitly expressed as the product of L -functions, Γ -functions, and exponential functions. This method of expressing L -functions by integrals on J_k and applying the Θ -formula can be applied to investigate Hey zeta functions and L -functions of various types defined for a simple algebra (\rightarrow

450 Zeta Functions) (G. Fujisaki [6]; T. Tamagawa, *Ann. Math.*, 77 (1963)).

References

- [1] J. Tate, Fourier analysis in number fields and Hecke's zeta-functions, Thesis, Princeton Univ., 1950. (Algebraic number theory, edited by J. W. S. Cassels and A. Fröhlich, Academic Press, 1967, ch. 15.)
- [2] A. Weil, Sur la théorie du corps de classes, *J. Math. Soc. Japan*, 3 (1951), 1–35.
- [3] E. Artin and J. Tate, Class field theory, Lecture notes, Harvard Univ., 1961 (Benjamin, 1967).
- [4] E. Artin, Representatives of the connected component of the idele class group, *Proc. Intern. Symp. Alg. Number Theory*, Tokyo and Nikko, 1955, 51–54.
- [5] G. Shimura and Y. Taniyama, Modern number theory (in Japanese), Kyôritu, 1957.
- [6] G. Fujisaki, On the zeta-function of the simple algebra over the field of rational numbers, *J. Fac. Sci. Univ. Tokyo*, 7 (1958), 567–604.
- [7] A. Weil, Adeles and algebraic groups, Lecture notes, Institute for Advanced Study, Princeton, 1961.
- [8] S. Lang, Algebraic numbers, Addison-Wesley, 1964.
- [9] A. Weil, Basic number theory, Springer, 1967.

7 (VI.15) Affine Geometry

A. Construction of Affine Spaces

An affine space A is constructed as follows: Let V be a \dagger vector space over a \dagger field K , and let A be a nonempty set. For any vector $\mathbf{a} \in V$ and any element p of A , suppose that an addition $p + \mathbf{a} \in A$ is defined so as to satisfy the following three conditions: (i) $p + \mathbf{0} = p$ ($\mathbf{0}$ being a zero vector); (ii) $(p + \mathbf{a}) + \mathbf{b} = p + (\mathbf{a} + \mathbf{b})$ ($\mathbf{a}, \mathbf{b} \in V$); and (iii) for any $q \in A$ there exists a unique vector $\mathbf{a} \in V$ such that $q = p + \mathbf{a}$. (Condition (i) follows from (ii) and (iii).) Then we call A an **affine space**, V the **standard vector space** of A , and K the **coefficient field** of A . Each element of A is called a **point**.

If we fix an arbitrary point $o \in A$, there is a one-to-one correspondence between A and V given by the mapping sending $p \in A$ to $\mathbf{a} \in V$ such that $p = o + \mathbf{a}$. Such an element \mathbf{a} of V is called a **position vector** of p with the **initial point** o and is denoted by \overrightarrow{op} . We say that $r + 1$

7 B

Affine Geometry

points p_α ($0 \leq \alpha \leq r$) of A are **independent** if r vectors $\mathbf{a}_i = \overrightarrow{p_0 p_i}$ ($1 \leq i \leq r$) are linearly independent in V ; otherwise, they are said to be **dependent**. This definition of dependence of points p_α is independent of the choice of the initial point among them. If V is of dimension n , we say that A is of **dimension** n , $\dim A = n$; in this case, we sometimes write A^n instead of A and V^n instead of V . The affine space A is of dimension n if and only if the maximum number of independent points in A is $n + 1$.

Next, for any vector subspace V^k of V^n and an arbitrary point $p \in A^n$, we put $A_p^k = \{q \in A^n \mid q = p + \mathbf{x}, \mathbf{x} \in V^k\}$ and call it a **subspace** of A^n . It is an affine space of dimension k : Conversely, every subset of A^n that is an affine space under the affine space structure of A can be expressed in this form. A^1 , A^2 , and A^{n-1} in A^n are called a **line**, **plane**, and **hyperplane**, respectively. A set that consists of only one point is also considered as a subspace A^0 . For subspaces A^r and A^s of A^n , we denote by $A^r \cap A^s$ the **intersection** (i.e., the set-theoretic intersection) of A^r and A^s , and by $A^r \cup A^s$ the **join** of A^r and A^s (i.e., the intersection of all subspaces that contain both A^r and A^s). Then $A^r \cap A^s$ is the affine space of highest dimension contained in A^r and A^s , and $A^r \cup A^s$ is the affine space of lowest dimension that contains A^r and A^s . If $r + 1$ points are given in A^n , there always exists a subspace A^r that contains all these points. In particular, if the points are independent, then such an A^r is unique. Moreover, if $A^r \cap A^s \neq \emptyset$ (\emptyset is the empty set), then we have $r + s = \dim(A^r \cup A^s) + \dim(A^r \cap A^s)$. This is called the **dimension theorem** (or **intersection theorem**) of affine geometry.

Next suppose that $r + 1$ points p_α ($0 \leq \alpha \leq r$) in A^n are independent, and put $p_{r+1} = p_0$. Let q_α be an arbitrary point on $p_\alpha \cup p_{\alpha+1}$ that differs from p_α and $p_{\alpha+1}$. If λ^α are elements of K such that $\lambda^\alpha \cdot \overrightarrow{p_\alpha q_\alpha} = \overrightarrow{q_\alpha p_{\alpha+1}}$, then q_0, \dots, q_r are dependent if and only if $\lambda^0 \lambda^1 \dots \lambda^r = (-1)^{r+1}$. And if $r \geq 2$ and $\sigma_\alpha = q_\alpha \cup p_{\alpha+2} \cup \dots \cup p_{\alpha+r}$ ($p_{i+r} = p_{i-1}$ if $i \geq 1$), then $\sigma_0, \dots, \sigma_r$ have a point in common if and only if $\lambda^0 \lambda^1 \dots \lambda^r = 1$. The former is called **Menelaus's theorem**, and the latter **Ceva's theorem**.

The set $L(A)$ of all subspaces (including \emptyset considered as an affine space of dimension -1) constitutes a \dagger lattice with respect to the inclusion relation.

B. Parallelism in Affine Spaces

Let A^r and A^s be subspaces of A^n . We say that A^r and A^s are **parallel in the wider sense** if either of the following conditions holds:

- (i) $A^r \supset A^s$ or $A^s \supset A^r$; or (ii) $A^r \cap A^s = \emptyset$ and $\dim(A^r \cup A^s) \leq r + s$. Next, let A^r and

B^r be subspaces of A^n of the same dimension. If A^r and B^r coincide, or $A^r \cap B^r = \emptyset$ and $\dim(A^r \cup B^r) = r + 1$, then they are said to be **parallel in the narrower sense** (or simply **parallel**), and we denote the relation by $A^r \parallel B^r$. If $r = s = 1$, the definitions of parallelism in the narrower sense and wider sense are equivalent. If $r > 1$, parallelism in the narrower sense implies parallelism in the wider sense. For two sets (a_α) and (b_α) ($0 \leq \alpha \leq r$) of $r + 1$ independent points, let V^r and W^r be vector spaces with bases $\overrightarrow{a_0 a_i}$ and $\overrightarrow{b_0 b_i}$ ($1 \leq i \leq r$), respectively. Then $A^r = a_0 \cup \dots \cup a_r$ and $B^r = b_0 \cup \dots \cup b_r$ are parallel if and only if $V^r = W^r$; and for an arbitrary point p , there exists a unique r -dimensional subspace that is parallel to A^r and passes through point p . If A^r and B^s are parallel in the wider sense, then there exist subspaces A^t and B^t ($t \geq 1$) of A^r and B^s that are parallel to each other. Moreover, if neither A^r nor B^s is contained in the other and if t is the largest integer with the property just given, then we have $t = r + s + 1 - \dim(A^r \cup B^s)$.

Parallelism between subspaces of A^n is an equivalence relation. Specifically, the equivalence class of a 1-dimensional subspace A^1 is called a **point at infinity** and is denoted by A_∞^0 . Given a subspace A^r of A^n , the set of points at infinity A_∞^0 represented by lines A^1 contained in A^r is denoted by A_∞^{r-1} ; we have $A^r \parallel B^r$ if and only if $A_\infty^{r-1} = B_\infty^{r-1}$. A_∞^{r-1} is called a **space at infinity**. In particular, the set A_∞^{n-1} is called the **hyperplane at infinity**. The set-theoretic sum $A^n \cup A_\infty^{n-1} = \bar{A}^n$ is supplied with the structure of a \dagger projective space; the "points" in \bar{A}^n are elements of \bar{A}^n , and the "lines" in \bar{A}^n are $A^1 \cup A_\infty^0$ and A_∞^1 .

C. Coordinates of Affine Spaces

If we fix a point o in A^n and a basis $\{e_1, \dots, e_n\}$ of the standard vector space V^n , then any point p in A^n is uniquely expressed as

$$p = o + \sum_{i=1}^n x^i \cdot e_i, \quad x^i \in K. \quad (1)$$

The system $\mathfrak{F} = (o; e_1, \dots, e_n)$ is called an **affine frame** (simply the **frame**) of A^n ; the point o is called its **origin**, and e_i is called the i th **unit vector**. The mapping sending p to (x^1, \dots, x^n) gives a \dagger bijection of A^n to K^n ; we call (x^1, \dots, x^n) **affine coordinates** of p with respect to \mathfrak{F} , and x^i the i th **affine coordinate**. In particular, if K is a topological field (e.g., the real number field \mathbf{R} or the complex number field \mathbf{C}), this bijection $A^n \rightarrow K^n$ induces a topology of A^n , which can be shown to be independent of the choice of \mathfrak{F} . For the remainder of this article, by "coordinates" we mean affine coordinates unless otherwise stated. Putting

$a_i = o + e_i$ ($1 \leq i \leq n$), we sometimes call $(o; a_1, \dots, a_n)$ an affine frame. Further, putting $l_i = o \cup a_i$, $\pi_i = o \cup a_1 \cup \dots \cup a_{i-1} \cup a_{i+1} \cup \dots \cup a_n$, we call a_i , l_i , and π_i the i th **unit point**, the i th **coordinate axis**, and the i th **coordinate hyperplane**, respectively.

Assume that subspaces A^r and A^s ($r, s > 0, r + s = n$) are not parallel in the wider sense. For a point p of A^n , denote by $A^r(p)$ the subspace that passes through p and is parallel to A^r , and put $q = A^r(p) \cap A^s$. A mapping $\varphi: A^n \rightarrow A^s$ defined by $\varphi(p) = q$ is called a **parallel projection** on A^s with respect to A^r . In particular, if $A^r = \pi_i$ and $A^s = l_i$ ($r = n - 1, s = 1$), we write $\varphi(p) = p_i$. Then the i th coordinate x^i of p is an element of K such that $\overrightarrow{op_i} = x^i \overrightarrow{oa_i}$. Hence such coordinates are also called **parallel coordinates** (or **Cartesian coordinates**).

Suppose that we are given $r + 1$ points b_0, \dots, b_r of A^n and $r + 1$ elements $\lambda^0, \dots, \lambda^r$ of K such that $\sum_{\alpha=0}^r \lambda^\alpha = 1$. We fix a point o of A^n . If a point p in A^n satisfies $\overrightarrow{op} = \sum_{\alpha=0}^r \lambda^\alpha \overrightarrow{ob_\alpha}$, then $\overrightarrow{b_0 p} = \sum_{i=1}^r \lambda^i \overrightarrow{b_0 b_i}$; hence p is contained in the subspace $b_0 \cup \dots \cup b_r$. Conversely, if a point p is contained in the latter subspace, then there exists a system $(\lambda^0, \dots, \lambda^r)$ such that

$$\overrightarrow{op} = \sum_{\alpha=0}^r \lambda^\alpha \overrightarrow{ob_\alpha} \quad \text{and} \quad \sum_{\alpha=0}^r \lambda^\alpha = 1.$$

The system $(\lambda^0, \dots, \lambda^r)$ has a geometric meaning since we also have $\overrightarrow{o'p} = \sum_{\alpha=0}^r \lambda^\alpha \overrightarrow{o'b_\alpha}$ if we replace the point o by any other point o' of A^n . The elements $\lambda^0, \dots, \lambda^r$ are called **barycentric coordinates** of p with respect to $\{b_0, \dots, b_r\}$. In particular, if $\{b_0, \dots, b_r\}$ are independent, then the barycentric coordinates $(\lambda^0, \dots, \lambda^r)$ are uniquely determined by the point p on $b_0 \cup \dots \cup b_r$. Furthermore, let (y^1, \dots, y^n) be affine coordinates of p with respect to an affine frame \mathfrak{F} , and let $(x_\alpha^1, \dots, x_\alpha^n)$ be affine coordinates of b_α ($\alpha = 0, \dots, r$). Then p belongs to the subspace $A^r = b_0 \cup \dots \cup b_r$ if and only if $y^i = \sum_{\alpha=0}^r \lambda^\alpha x_\alpha^i$ ($i = 1, \dots, n$). In this case we say that the system of the linear equations $y^i = \sum \lambda^\alpha x_\alpha^i$ ($\sum \lambda^\alpha = 1$) gives a **parametric representation** of the subspace A^r (by parameters λ^α). Specifically, if $r = n - 1$, the solvability of the system of equations $y^i = \sum_{\alpha=0}^{n-1} \lambda^\alpha x_\alpha^i$ ($i = 1, \dots, n$), $1 = \sum \lambda^\alpha$ implies the equation $\sum_{i=1}^n y^i \mu_i = \mu_0$ for some nontrivial constants μ_0, \dots, μ_n . Hence the latter equation represents the hyperplane $\pi = A^{n-1}$. If a point p has barycentric coordinates $\lambda^0 = \lambda^1 = \dots = \lambda^r = (r + 1)^{-1}$ with respect to $\{b_0, \dots, b_r\}$, it is called the **barycenter** of b_0, b_1, \dots, b_r . The barycenter is uniquely determined by the set $\{b_0, \dots, b_r\}$ and is denoted by $g(b_0, \dots, b_r)$. Specifically, the barycenter of two points b_0 and b_1 is called the **midpoint** (or **middle point**) of b_0 and b_1 . If we divide $B = \{b_0, \dots, b_r\}$ into two sets of points and if g_1 and g_2 are barycenters of these two sets of points, respectively, then $g_1 \cup g_2$ passes

through the barycenter of B . More generally, a point with barycentric coordinates $(\lambda^0, \dots, \lambda^r)$ with respect to $\{b_0, \dots, b_r\}$ is called a barycenter of b_0, \dots, b_r with **weights** $\lambda^0, \dots, \lambda^r$.

D. Affine Spaces over Ordered Fields

Suppose the coefficient field K is an \dagger ordered field (e.g., the real number field \mathbf{R}). Given a hyperplane π of A^n , we take an affine frame for which π is the n th coordinate hyperplane. If we denote coordinates of points with respect to this frame by (x^1, \dots, x^n) , then the equation of π is given by $x^n = 0$. Let A_+^n and A_-^n be sets of points whose n th coordinates are positive and negative, respectively. They are called **half-spaces** of A^n divided by π . The union of π and a half-space is called a **closed half-space**. A half-space of a subspace A^r of A^n (divided by some A^{r-1} on A^r) is called a half-space of dimension r . For a point p of A^n that does not lie on π , the half-space containing p is called the **side** of p with respect to π . In particular, when $n = 1$, let p and q be two points on a line l . The closed side of q with respect to p is called the **(closed) half-line** (or **ray**) from p to q . The intersection of the closed half-lines emanating from p to q and from q to p is called the **segment** joining p and q and is denoted by \overline{pq} . Clearly $\overline{pq} = \overline{qp}$. A subset C of A^n is called a **convex set** if the segment joining two arbitrary points of C is contained in C . Each half-space of each dimension is convex. For any family C_i of convex sets, $\bigcap_i C_i$ is also convex. Therefore, for any subset D in A^n there exists a minimal convex set that contains D . It is called the **convex closure** (or **convex hull**) of D . The convex closure $C(P)$ of a finite set of points $P = \{p_0, \dots, p_k\}$ in A^n is called a **convex cell**, and $\dim(p_0 \cup \dots \cup p_k)$ is called the **dimension** of the convex cell. In particular, when p_0, \dots, p_k are independent, $C(P)$ is called a k -dimensional **simplex** with **vertices** p_0, \dots, p_k . The 1-dimensional simplex having two distinct points p and q as vertices is the segment \overline{pq} , and the vertices p and q are called **ends** of the segment. A point is regarded as a 0-dimensional simplex. Each 2-dimensional or 3-dimensional simplex is called a **triangle** or **tetrahedron**, respectively. A k -dimensional simplex S with vertices p_0, \dots, p_k is a set of points whose barycentric coordinates λ^α ($0 \leq \alpha \leq k, \sum \lambda^\alpha = 1$) with respect to the vertices satisfy $\lambda^\alpha \geq 0$. On the other hand, if we put $A^k = p_0 \cup \dots \cup p_k$ and $\pi_\alpha = p_0 \cup \dots \cup p_{\alpha-1} \cup p_{\alpha+1} \cup \dots \cup p_k$, and denote by A_α^k the side of p_α in A^k and by \bar{A}_α^k the closed side of p_α in A^k with respect to π_α , then the simplex S is given by $\bigcap_{\alpha=0}^k \bar{A}_\alpha^k$, and $\bigcap_{\alpha=0}^k A_\alpha^k$ is called an **open simplex**. A^n has the structure of a \dagger topological space in which the set of open n -

dimensional simplexes forms a \dagger base of \dagger open sets. In particular, if K is \mathbf{R} , the topology of A^n thus defined is compatible with the one that is naturally induced by the topology of \mathbf{R} . With respect to this topology, A^n is a \dagger Hausdorff space. The terms *open* and *closed* used before for n -dimensional simplexes agree with the corresponding notions with respect to this topology.

A subset of A^n is said to be **bounded** if it is contained in some simplex. A bounded set obtained through a finite process of constructing intersections and unions from a finite number of closed half-spaces is called a **polyhedron**. The points of a convex polyhedron are characterized by several linear inequalities satisfied by their coordinates. A set of points whose coordinates (x^1, \dots, x^n) satisfy $h^i \leq x^i \leq k^i$ for $k^i, h^i \in K$ is called a **parallelootope**; it is a polyhedron whose \dagger interior is called an **open parallelootope**. A simplex is a polyhedron, and polyhedra admit \dagger simplicial decompositions. A polyhedron can also be defined as the set-theoretic union of a finite number of simplexes.

Let P be a finite set of points, and let its convex closure $C(P)$ be a convex cell of dimension m . Then $C(P)$ is contained in an m -dimensional subspace A^m . The \dagger boundary of $C(P)$ in the topological space A^m is called the **boundary** of a convex cell $C(P)$. We can take a subset Q of P so that $\dim C(Q) = m - 1$ and $C(Q)$ is the intersection of the boundary of $C(P)$ and an $(m - 1)$ -dimensional subspace. Such a $C(Q)$ is called a **face** of $C(P)$, and we denote this relation by $C(P) > C(Q)$. If $C(P) > C(P_1) > \dots > C(P_s)$, then $C(P_s)$ is called an $(m - s)$ -dimensional face of $C(P)$. A 0-dimensional face is called a **vertex**, and a 1-dimensional face is called an **edge**. Suppose that $C(P) > C(Q)$ for $P = \{p_0, \dots, p_k\}$ and $Q = \{p_{i_0}, \dots, p_{i_{k-1}}\}$. Then $F = p_{i_0} \cup \dots \cup p_{i_{k-1}}$ is a hyperplane of $E = p_0 \cup \dots \cup p_k$, and $C(P)$ is contained in a closed side of E divided by F . Therefore, if $C(P)$ has d $(m - 1)$ -dimensional faces, then $C(P)$ is expressed as the intersection of d m -dimensional closed half-spaces. This shows that any convex cell is a polyhedron.

E. Affine Transformations

A mapping $\varphi: A^n \rightarrow A^m$ is an **affine mapping** if there is a linear mapping $\bar{\varphi}: V^n \rightarrow V^m$ of the standard vector spaces of A^n and A^m such that $\varphi(p + \mathbf{x}) = \varphi(p) + \bar{\varphi}(\mathbf{x})$ holds for any $p \in A^n$ and any $\mathbf{x} \in V^n$. An affine mapping of A^n into itself is called an **affine transformation** (or **affinity**) of A^n . Specifically, a bijective affine transformation is called a **regular** (or **proper**) **affine transformation**. An affine transformation φ of A^n is characterized by each one of the following

properties: (i) Let $o \in A''$ be a fixed point. Then φ is a mapping of A'' onto itself that can be expressed as

$$\varphi(o + \mathbf{x}) = o + \mathbf{a} + f(\mathbf{x}), \quad (2)$$

where \mathbf{a} is a fixed vector of V'' and f is a linear transformation of V'' . (ii) The mapping $\varphi: A'' \rightarrow A''$ is a mapping such that $\varphi(a)\varphi(b) = \lambda \cdot \varphi(p)\varphi(q)$ if $\overline{ab} = \lambda \cdot \overline{pq}$ ($\lambda \in K$). Moreover, if the characteristic of K is not equal to 2, an affine transformation is also characterized as follows: (iii) φ is a mapping that sends lines into lines and preserves the ratio of each pair of parallel segments.

The set $\mathfrak{U}(A'')$ of all regular affine transformations of A'' constitutes a group that we call the **group of affine transformations**. If the linear mapping f associated with a regular affine transformation φ is the identity mapping, then φ is called a **translation**. The set $\mathcal{T}(A'')$ of all translations is a normal subgroup of $\mathfrak{U}(A'')$ and is called the **group of translations**. The group of translations is isomorphic to V regarded as an additive group. The vector group $\mathfrak{B}(A'')$ (i.e., an additive group of a linear space) acts simply transitively on A'' . We see that $\mathfrak{U}(A'')/\mathfrak{B}(A'') \cong GL(n, K)$, where $GL(n, K)$ denotes the general linear group. The set of all regular affine transformations that leave a point o of A'' invariant constitutes a subgroup $\mathfrak{G}(A'')$ of $\mathfrak{U}(A'')$; it is called an **isotropy group** at o and is isomorphic to $GL(n, K)$. Let $\mathfrak{F} = (o; \mathbf{e}_1, \dots, \mathbf{e}_n)$ be an affine frame of A'' with origin o , and let φ be a regular affine transformation of A'' given by (2); put $\mathbf{x} = \sum x^i \mathbf{e}_i$, $\varphi(o + \mathbf{x}) = o + \sum \bar{x}^i \mathbf{e}_i$, $\mathbf{a} = \sum a^i \mathbf{e}_i$, and $f(\mathbf{e}_i) = \sum a_j^i \mathbf{e}_j$. Then φ is expressed with respect to \mathfrak{F} by the following equation:

$$\bar{x}^i = a^i + \sum_{k=1}^n a_k^i x^k, \quad \det(a_j^i) \neq 0, \quad 1 \leq i \leq n. \quad (3)$$

Conversely, a transformation that is given by (3) is a regular affine transformation. Elements of $\mathfrak{B}(A'')$ and $\mathfrak{G}(A'')$ are expressed with respect to \mathfrak{F} by

$$\bar{x}^i = x^i + a^i, \quad 1 \leq i \leq n, \quad (4)$$

and

$$\bar{x}^i = \sum_{k=1}^n a_k^i x^k, \quad \det(a_j^i) \neq 0, \quad 1 \leq i \leq n, \quad (5)$$

respectively. Hence $\mathfrak{U}(A'')$ is represented as a **semidirect product** group of $\mathfrak{B}(A'')$ and $(F)(A'')$. In particular, a regular affine transformation that is represented by $\bar{x}^i = ax^i$ ($1 \leq i \leq n$) for some $a \in K$ ($a \neq 0$) is called a **similarity** with the origin o as center.

According to F. Klein, the objects we deal with in affine geometry are the properties (parallelism, barycenters, etc.) that are invariant under regular affine transformations.

Subsets S_1 and S_2 of A'' are called **affinely congruent** if there exists a regular affine transformation φ sending S_1 onto S_2 . For a fixed k , two k -dimensional simplexes are affinely congruent. Now we fix an affine frame \mathfrak{F} in A'' and denote by x_a^i the coordinates of $n+1$ points p_a ($0 \leq a \leq n$) in A'' . Then the quantity

$$V(p_0, \dots, p_n) = \frac{1}{n!} \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_0^1 & x_1^1 & \dots & x_n^1 \\ \dots & \dots & \dots & \dots \\ x_0^n & x_1^n & \dots & x_n^n \end{vmatrix} \quad (6)$$

is called the **volume** with respect to \mathfrak{F} of the n -dimensional simplex with vertices p_0, \dots, p_n . If φ is a regular affine transformation given by (3), we have $\bar{V} = \det(a_j^i) V$. Hence the ratio of volumes of two n -dimensional simplexes is independent of the choice of coordinate systems, and is invariant under regular affine transformations.

A regular affine transformation given by (3) satisfying $\det(a_j^i) = 1$ is called an **equivalent affinity**. The set of all equivalent affinities constitutes a subgroup of $\mathfrak{U}(A'')$. The geometry belonging to this group is called **affine geometry in the narrower sense**. For instance, the concept of volume is an invariant in affine geometry in the narrower sense.

F. Relation to Projective Geometry

Let \mathbf{P}^n be a projective space over a coefficient field K (\rightarrow 343 Projective Geometry). If we fix a hyperplane π_∞ in \mathbf{P}^n , then the set of projective transformations that leave π_∞ invariant constitutes a subgroup of the group of projective transformations of \mathbf{P}^n ; this subgroup is isomorphic to a group of regular affine transformations. Actually, if we use a projective frame $[a_0, a_1, \dots, a_n, u]$ such that a_0, \dots, a_n are points on π_∞ , then each projective transformation leaving π_∞ invariant is expressed by equations of the same form as (3) with respect to the inhomogeneous projective coordinates. The point set A'' complementary to π_∞ in \mathbf{P}^n is an affine space, and π_∞ coincides with the hyperplane at infinity. Moreover, two distinct lines in \mathbf{P}^n are parallel in A'' if they meet on the hyperplane at infinity. Hence, denoting by $(0, l^1, \dots, l^n)$ the homogeneous projective coordinates of the intersection of a line l in A'' and π_∞ , we call (l^1, \dots, l^n) the **direction ratio** of l . A projective transformation leaving each point of π_∞ invariant induces a translation. The principle of duality that holds in projective geometry does not hold in affine geometry. The pole of the hyperplane at infinity, with respect to a quadric hypersurface, is called the **center** of the quadric hypersurface. A regular quadric hypersurface is called **central**

or **noncentral** according as its **center** belongs to A^n or is a point at infinity. Quadric hypersurfaces in an affine space are classified in several ways, by taking account of their relations with the hyperplane at infinity (→ 78 Conic Sections, 350 Quadric Surfaces).

References

- [1] O. Schreier and E. Sperner, Einführung in die analytische Geometrie und Algebra I, II, Teubner, 1931, 1935; English translation, Introduction to modern algebra and matrix theory, Chelsea, second edition, 1961.
- [2] E. Sperner, Einführung in die analytische Geometrie und Algebra I, II, Vandenhoeck & Ruprecht, 1948, 1951.
- [3] E. Artin, Geometric algebra, Interscience, 1957.
- [4] H. Weyl, Raum, Zeit, Materie, Springer, fifth edition, 1923; English translation, Space, time, matter, Dover, 1952.
- [5] S. Iyanaga and K. Matsuzaka, Affine geometry and projective geometry, J. Fac. Sci. Univ. Tokyo, 14 (1967), 171-196.

8 (111.1) Algebra

Though the word “algebra” usually refers to a field of mathematics as will be explained below, the word may also denote specific mathematical structures such as \dagger associative algebras, \dagger Jordan algebras, \dagger Clifford algebras, etc. The first concepts concerning “unknowns” in algebra originated in India, whence came also our decimal positional system of numeration. These ideas were introduced to Europe through Arabia in the Renaissance period. F. \dagger Viète systematized them into a symbolic method, called **algebra**, representing numbers by letters. The first problem of algebra was solving equations. Before Viète, G. Cardano and L. Ferrari had solved algebraic equations of degrees 3 and 4; the solution of equations of lower degree had been known from antiquity. The effort to solve equations of higher degree remained unresolved until the middle of the 19th Century, when N. H. \dagger Abel and E. \dagger Galois proved the nonexistence of algebraic solutions of such equations. They considered not only individual roots of these equations but also any rational transforms of their roots at the same time, and thus were led to the concept of \dagger fields. They also noticed that the problem of algebraic solution could be characterized by properties of permutation groups of the roots.

After the discovery of the Galois group, group theory and group-theoretical considerations maintained the central position in algebra for some time (→ 172 Galois Theory). These developed into the “abstract algebra” of this Century in the general atmosphere of arithmetization and of axiomatization of mathematics. At the turn of the century the monumental textbook in three volumes by H. Weber [1] was considered the standard work on algebra. Then there appeared in 1910 an epoch-making paper [2] by E. Steinitz on the abstract theory of fields.

The main objects of algebra today are \dagger algebraic systems of various kinds, such as \dagger groups, \dagger rings, \dagger fields, and \dagger modules. Another fundamental concept of algebra is that of \dagger isomorphism or of \dagger homomorphism. The collection of algebraic systems of a given kind, together with the homomorphisms among them, gives rise to the notion of \dagger category; a functor is a sort of homomorphism between categories (→ 52 Categories and Functors). These notions were first used in \dagger homological algebra, created in the 1940s by methods transferred from topology to algebra; now they are of basic significance to the whole of mathematics.

An important branch of algebra with wide applications is the theory of \dagger vector spaces, or more generally that of \dagger modules over a ring. This branch is called **linear algebra**. Homomorphisms between finitely generated modules can be represented by \dagger matrices. Another branch of algebra, called \dagger representation theory, is concerned with representations of groups or rings by matrices. The methods of modern algebra provide useful and powerful tools for the whole of mathematics, in particular for the theory of numbers and algebraic geometry.

The present development of algebra owes much to the activity of the German school in the late 1920s represented by E. Noether, E. Artin, W. Krull, and B. L. van der Waerden. The book by van der Waerden [3] has had a great impact on mathematics. N. Bourbaki [4] has been influenced by van der Waerden but gives accounts of more recent developments, particularly in linear algebra. In Japan, M. Sono, who worked at about the same period as E. Noether, was a forerunner in this field; after him, algebraists of the Kyoto School, Y. Akizuki, M. Nagata, and their followers, did notable research, especially in algebraic geometry. On the other hand, K. Shoda studied with E. Noether toward 1930 in Germany; his school includes such algebraists as T. Nakayama, K. Asano, and G. Azumaya. Finally K. Morita and his disciples have made significant contributions to homological algebra.

References

- [1] H. M. Weber, *Lehrbuch der Algebra*, Vieweg, I, 1894; II, 1896; III, 1891.
 [2] E. Steinitz, *Algebraische Theorie der Körper*, J. Reine Angew. Math., 137 (1910), 167–309.
 [3] B. L. van der Waerden, *Moderne Algebra*, Springer, first edition, I, 1930; II, 1931; English translation, *Algebra I, II*, Ungar, 1970.
 [4] N. Bourbaki, *Eléments de mathématique II. Algèbre*, ch. I-9, *Actualités Sci. Ind.*, 1144b, 1236b, 1044, 1102b, 1179a, 1261a, 1272a, Hermann, 1958-1964.
 [5] G. Birkhoff and S. MacLane, *A survey of modern algebra*, Macmillan, third edition, 1965.

9 (VIII.2) Algebraic Curves

A. General Remarks

An \dagger algebraic variety of dimension 1 is called an **algebraic curve** (for analytic theory \rightarrow 11 Algebraic Functions). The theory of algebraic curves has two aspects, the geometry of 1-dimensional complex manifolds in projective spaces and the theory of function fields of transcendence degree 1 (\rightarrow 3 Abelian Varieties, 16 Algebraic Varieties). The number-theoretic study of algebraic function fields concerns the latter theory (\rightarrow 73 Complex Multiplication, 450 Zeta Functions). In this article, the geometric aspect of the theory is emphasized. We denote the \dagger universal domain by **K**.

B. Classical Results on Plane Algebraic Curves

Let $f(X, Y)$ be a polynomial of degree m in two variables X and Y . A point set in an affine two-space defined by $f(X, Y) = 0$ is called the plane algebraic curve of degree m . If we set $F(X_0, X_1, X_2) = X_0^m f(X_1/X_0, X_2/X_0)$, the homogeneous polynomial F defines an algebraic curve of degree m in a projective plane P^2 . The curve is called **irreducible** if the polynomial $f(X, Y)$ is irreducible. A curve of degree 1 is said to be a line. Some results of this section are valid only in the case where the characteristic of **K** is zero.

Let C be a plane curve defined by the equation $f(X, Y) = 0$. A point $P = (a, b)$ on C is called an **r-ple point** if $f(X + a, Y + b)$ has no term of degree $< r$ in X and Y . At an r -ple point there are r tangent straight lines (counting multiplicity). An r -ple point with $r > 1$ is

called a **multiple point** (or **singular point**). If these tangents are distinct, P is called **ordinary**. An ordinary double point is called a **node**; e.g., the origin for $X^3 + Y^3 - 3XY = 0$. An algebraic curve can be transformed into a plane curve that has only ordinary multiple points by a finite number of plane \dagger Cremona transformations (\dagger quadratic transformations of the projective plane into itself).

Let C be an irreducible plane curve of degree $m > 1$ in a projective plane. The set of all tangent lines at nonsingular points of C determines a set of points in the dual projective plane, and its closure is an algebraic curve \hat{C} , that is called the **dual curve** of C . The dual curve of \hat{C} becomes the original C . The degree m' of \hat{C} is said to be the **class** of C , which is equal to the number of tangent lines to C drawn from a general point. A nonsingular point P is called a **point of inflection** (or a **flex**) if the tangent line at P has contact of order > 2 . If C is defined by an irreducible homogeneous polynomial $F(X_0, X_1, X_2)$, the curve defined by $\det(\partial^2 F / \partial X_i \partial X_j) = 0$ is said to be the **Hessian** of C . A nonsingular point P is a point of inflection if and only if P is contained in the Hessian. A singular point P is said to be a **cusp** if C is defined by an equation $Y^2 = X^3 + \text{higher terms}$, in terms of suitable affine coordinates X and Y where $P = (0, 0)$. Whenever the singular points of C are only v nodes and y cusps, the effective genus g is given by the formula $g = (m-1)(m-2)/2 - v - y$. In addition, suppose that the dual curve has only v' nodes and y' cusps as singularities. Then C has y' points of inflection, and m' and y' are given by $m' = m(m-1) - 2v - 3y$ and $y' = 3m(m-2) - 6v - 8y$. Moreover, $m = m'(m'-1) - 2v' - 3y'$ and $y = 3m'(m'-2) - 6v' - 8y'$ hold. These formulas are called **Plücker formulas**.

For example, a nonsingular plane curve of degree 3 is an elliptic curve, i.e., $g = 1$, of class 6, which has 9 points of inflection. A nonsingular plane curve of degree 4 has, in general, 24 points of inflection and 28 bitangents. This results from the Plücker formulas.

C. Fundamental Notions

In what follows, by a curve we mean an algebraic variety of dimension 1. Let Γ be a nonsingular complete irreducible curve. An element of the free Abelian group generated by points of Γ is called a **divisor**. A divisor is written in the form $a = \sum n_i P_i$, with $n_i \in \mathbb{Z}$. The integer $n = \sum n_i$ is called the **degree** of a and is denoted by $\deg a$. The expression for a divisor a is said to be reduced if $P_i \neq P_j$ for $i \neq j$. A divisor whose reduced expression has only positive coefficients is called a **positive divisor** (or **effective**).

tive divisor, or integral divisor), and this is denoted by $\alpha \succ 0$. The group of divisors on Γ is denoted by $G(\Gamma)$, and the subgroup consisting of divisors of degree 0 is denoted by $G_0(\Gamma)$. Let P be a point of Γ . The subset of the function field $\mathbf{K}(\Gamma)$ of Γ consisting of functions regular at P forms a valuation ring R_P for a †discrete valuation of $\mathbf{K}(\Gamma)$. A prime element t of R_P is called a **local parameter** at P . Let v_P be the †normalized valuation of $\mathbf{K}(\Gamma)$ defined by R_P ; the integer $v_P(f)$ is called the **order** off at P . The point P is a **zero** off if $v_P(f) > 0$; it is a **pole** off if $v_P(f) < 0$. There are only a finite number of poles and zeros of a given function f . The divisor $\sum v_P(f)P$ is called the **divisor of the function** f and is denoted by (f) . The set of divisors of functions forms a subgroup G_f of G . Any divisor α in G_f is called a **principal divisor** (we also say that α is †linearly equivalent to zero and write $\alpha \sim 0$).

Let α be an arbitrary divisor. The set of all positive divisors that are linearly equivalent to α forms a **complete linear system** $|\alpha|$ determined by α . We set $L(\alpha) = \{f \in \mathbf{K}(\Gamma) \mid (f) + \alpha \succ 0\} \cup \{0\}$. Then $L(\alpha)$ is a finite-dimensional vector space over \mathbf{K} , and 1-dimensional subspaces of $L(\alpha)$ correspond bijectively to the elements of $|\alpha|$. We set $J(\alpha) = \dim L(\alpha)$ and $\dim |\alpha| = J(\alpha) - 1$. Then $\dim |\alpha|$ is called the **dimension** of $|\alpha|$. For any divisor α , the integer $\deg \alpha = \dim |\alpha|$ is nonnegative and bounded. The supremum g of such integers is called the **genus** of Γ . The nonnegative integer $i(\alpha) = g - \deg \alpha + \dim |\alpha|$ is called the **speciality index** of α .

Let ω be a †differential form on Γ , P be a point of Γ , and t be a local parameter at P . Then ω can be written in the form $\omega = f dt$. We now set $v_P(\omega) = v_P(f)$ and $(\omega) = \sum v_P(\omega)P$. Then (ω) is a well-defined divisor, and the class of (ω) in G/G_f is independent of the choice of ω . This divisor class is called the **canonical class**; any divisor in this class is called a **canonical divisor** (or **differential divisor**) and is denoted by \mathfrak{f} . We have $J(\mathfrak{f}) = g$, $\deg \mathfrak{f} = 2g - 2$. Given a divisor α , the index $i(\alpha)$ is equal to the number of linearly independent differentials ω such that $(\omega) \succ \alpha$, $i(\alpha) = J(\mathfrak{f} - \alpha)$. The equality $J(\alpha) = \deg \alpha - g + 1 + i(\alpha)$ is called the **Riemann-Roch theorem**.

For any irreducible algebraic curve Γ , there exists a birationally equivalent nonsingular curve $\tilde{\Gamma}$ that is unique up to isomorphism. The genus of $\tilde{\Gamma}$ is called the **effective genus** of Γ . A curve whose effective genus is zero is called a **rational** (or **unicursal**) curve. An **elliptic curve** is a curve whose effective genus is 1. An irreducible curve Γ with genus ≥ 2 is called a **hyperelliptic curve** if $\mathbf{K}(\Gamma)$ is a quadratic extension of a field $\mathbf{K}(t)$ for some t .

A positive divisor α on a nonsingular com-

plete irreducible curve Γ is called a **special divisor** if $i(\alpha) > 0$. In this case, $2(J(\alpha) - 1) \leq \deg \alpha$, where equality holds if and only if $\alpha = 0$ or $\alpha \sim \mathfrak{f}$, or Γ is a hyperelliptic curve (Clifford's theorem). Let Γ be a nonsingular complete irreducible curve and k be a subfield of \mathbf{K} such that Γ is defined over k . Denoting by \bar{k} the algebraic closure of k , we call a divisor $p = \sum n_i P_i$ on Γ a **prime rational divisor over k** if p satisfies the following three conditions: (i) p is invariant under any automorphism σ of \bar{k}/k ; (ii) for any j , there exists an automorphism σ_j of \bar{k}/k such that $P_j = P_i^{\sigma_j}$; (iii) $n_1 = \dots = n_l = [k(P_1):k]$. An element in the subgroup of $G(\Gamma)$ generated by prime rational k -divisors is called a **k -rational divisor**. Let $k(\Gamma)$ be the subset of $\mathbf{K}(\Gamma)$ consisting of functions f defined over k . Then $k(\Gamma)$ is a subfield of $\mathbf{K}(\Gamma)$, and the quotient field of $k(\Gamma) \otimes_k \mathbf{K} = \mathbf{K}(\Gamma)$. $k(\Gamma)$ is called the **function field of Γ over k** . Let p be a prime rational k -divisor, and let P be a point of p . Then $R_P \cap k(\Gamma)$ is a valuation ring of $k(\Gamma)$ uniquely determined by p and independent of the choice of the point P in p . We call this valuation ring the valuation ring determined by p .

D. Algebraic Function Fields

Let k be a field, and let K be a finite separable extension of a purely transcendental extension $k(x)$ of k such that k is maximally algebraic in K . Then K is called an **algebraic function field over k of dimension 1** (or of **transcendence degree 1**). The equivalence class of †exponential valuations of K that are trivial over k is called a **prime divisor** of K/k . An element of the free Abelian group generated by prime divisors is called a **divisor** of K/k . The group operation in the divisor group of K/k is usually denoted multiplicatively. Let R_p be the valuation ring of the prime divisor P , and let M_p be the maximal ideal of R_p . The **degree** $\deg P$ of the prime divisor P is defined by $[(R_p/M_p):k]$. If we replace the terms: curve Γ by function field K/k ; $\mathbf{K}(\Gamma)$ by K ; \mathbf{K} by k ; and points on Γ by prime divisors of K/k , we can develop the theory of the function field K/k , which is similar to the theory of nonsingular curves Γ (Sections B, C). Thus we define the **genus of the function field K/k** .

Suppose we are given an algebraic function field K/k of dimension 1. An algebraic curve Γ defined over k is called a **model** of K/k if $k(\Gamma)$ and K are k -isomorphic. For any function field of dimension 1, there always exist two elements x and y in K such that $K = k(x, y)$. Let $f(X, Y)$ be an irreducible polynomial such that $f(x, y) = 0$. Then the plane curve defined by the equation $f(X, Y) = 0$ is a model of K/k . Among

the models of K/k there exists a †normal model Γ_0 over k that is unique up to isomorphism (and the uniqueness of the normal model of the function field within the birational equivalence class of varieties holds only for curves). In particular, if k is the complex number field, the normal model Γ_0 is the †Riemann surface of the function field K/k . If Γ_0 has no singular point, the theory of the curve Γ_0 and the theory of the function field K/k are essentially identical. (This occurs, for example, when k is †perfect.) In that case the genus of Γ_0 is equal to that of K/k . In general, the genus of the function field is not less than the genus of the normal model Γ_0 , and it is greater than the latter if Γ_0 has a singular point. If the genus of K/k is zero, we can take a plane quadratic curve as a model of K/k . Moreover, K/k has a prime divisor of degree 1 if and only if K is a purely transcendental extension of k . A function field K/k of genus 1 is called an **elliptic function field**. If K has a prime divisor of degree 1, an elliptic function field K has a model of a plane cubic curve. Moreover, if the characteristic of the field k is different from 2, we can take as the model Γ_0 the curve defined by an equation of the form $Y^2 = 4X^3 - g_2X - g_3$. This is called **Weierstrass's canonical form**. The number $j = (g_2^3 - 27g_3^2)^{-1}g_2^3$ (#0) is a birational invariant of Γ_0 .

E. Jacobian Varieties

Let Γ be a nonsingular curve. A †group variety J is called the **Jacobian variety** of Γ if it has the following four properties (we fix an algebraically closed †field k of definition for Γ and J): (i) There exists an isomorphism Φ (of abstract groups) of $G_0(\Gamma)/G_1(\Gamma)$ into J . (ii) Φ is continuous in the following sense: Let \bar{a}, \bar{b} be elements of $G_0(\Gamma)/G_1(\Gamma)$ represented by a, b . If b is a specialization of a over a field K ($\supset k$), then $\Phi(b)$ is also a specialization of $\Phi(a)$ over K . (iii) If there exists a K -rational divisor in the class \bar{a} , then the point $\Phi(\bar{a})$ is also K -rational. (iv) For any $\xi \in J$, there exists a $k(\xi)$ -rational divisor α in G_0 such that $\Phi(\alpha \bmod G_1) = \xi$. A group variety J satisfying these conditions is necessarily a complete variety, hence a †Abelian variety, and is determined uniquely up to isomorphism. The construction of Jacobian varieties over a field of arbitrary characteristic is due to A. Weil [27] (for analytic construction \rightarrow 11 Algebraic Functions C).

Let P be a †generic point of Γ over k , and let P_0 be a fixed k -rational point. Then $\varphi(P) = \Phi(P - P_0)$ defines a rational mapping of Γ into J , and φ , which is an isomorphism of Γ and its image $\varphi(\Gamma)$, is determined uniquely by Φ up to translation on J . This mapping φ is called the

canonical function on Γ . The dimension of J is equal to the genus g of Γ . If P_1, \dots, P_g are independent generic points of Γ over k , then $k(P_1, \dots, P_g)$ is the function field of J over k , where $k(P_1, \dots, P_g)$ is the subfield invariant under the group of $g!$ automorphisms $(P_1, \dots, P_g) \rightarrow (P_{\alpha_1}, \dots, P_{\alpha_g})$. The Jacobian variety of Γ is also the †Picard variety of Γ , and it is equal to the †Albanese variety of Γ (- 16 Algebraic Varieties P). Hence for any function f on Γ with values in an Abelian variety A , there exists a unique homomorphism λ of J into A such that $f = \lambda \circ \varphi + \text{const}$. This λ is called the **linear extension** of f .

Let Θ be the set of points on J that can be written as $\varphi(P_1) + \dots + \varphi(P_{g-1})$. Then Θ is an irreducible subvariety of codimension 1. The divisor Θ is called the **canonical divisor** of J . The Jacobian variety that is polarized by the divisor Θ is called the **canonically polarized Jacobian variety** (\rightarrow 3 Abelian Varieties G). If two curves Γ and Γ' are birationally equivalent, the canonically polarized Jacobians of Γ and Γ' are isomorphic. Conversely, if the canonically polarized Jacobian varieties J of Γ and J' of Γ' are isomorphic, then Γ and Γ' are birationally equivalent (**Torelli's theorem**). Let r be any integer such that $1 \leq r \leq g$, and let W_r be the set of points that are written in the form

$$\varphi(P_1) + \dots + \varphi(P_r)$$

($W_1 = \varphi(\Gamma)$, $W_{g-1} = \Theta$, $W_g = J$). Then we have $\Theta^{(r)} = r! W_{g-r}$ (†numerically equivalent) and $(\Theta^{(g)}) = y!$, where $\Theta^{(r)}$ is the class of intersections of r copies of Θ . The existence of a divisor Θ is characteristic for Jacobian varieties. Actually, if A is an Abelian variety of dimension n that has an irreducible subvariety X^{n-1} of codimension 1 and a positive 1-cycle C such that $(X^{(n)}) = n!$ and $X^{(n-1)} = (n-1)! C$, then C is a nonsingular irreducible curve, A is the Jacobian variety of C , and X is the canonical divisor. The canonical divisor Θ is defined by a †theta function in the classical case. For a divisor $\alpha = \sum n_i P_i$, we define $\varphi(\alpha)$ to be $\sum n_i \varphi(P_i) \in J$. For fixed numbers $1 \leq r \leq g$, $0 < d$, we let W_r^d denote $\{\varphi(\alpha) \in J \mid l(\alpha) \geq d+1, \deg \alpha = r\}$. Then $W_r^d \subseteq W_r$ and is a †subscheme of J . We call the number $\rho = g - (d+1)(g-r+d)$ the **Brill-Noether number**. Then $\dim W_r^d \geq \rho$. Moreover, if Γ is a general curve, then $\dim W_r^d = \rho$. In particular, if $\rho < 0$, then $W_r^d = \emptyset$. This result has been verified recently by S. Kleiman, D. Laksov, P. Griffiths, and J. Harris [9, 14].

Let Γ be a nonsingular curve, and let ω be a differential form on Γ . If the divisor (ω) is ≥ 0 , the ω is called a regular 1-form or **differential form of the first kind**. Let Ω be the †sheaf of germs of regular differential forms. A differential form of the first kind is an element of

$H^0(\Gamma, \Omega)$, and vice versa. Let \mathfrak{f} be a canonical divisor. Then we have a natural isomorphism $H^0(\Gamma, \Omega) \cong L(\mathfrak{f})$, and the number of linearly independent differential forms of the first kind is equal to the genus g of Γ . The \dagger residue of a differential can be defined as in the classical case. A differential that has nonzero residues is called a **differential of the third kind**. The **residue theorem** $\sum \text{Res}_P \omega = 0$ holds for any differential ω . The form ω is called a **differential form of the second kind** if for any $P \in \Gamma$ there exists a rational function f_P such that $\omega - df_P$ is regular at P . The set of differential forms of the second kind forms a linear space G_2 over the universal domain and contains the subspace G , consisting of the differential forms of the first kind. The quotient space $G_2/dk(\Gamma)$ has dimension 29 or y according as the characteristic of the universal domain is 0 or not.

When the characteristic p of the universal domain is positive, we have what is called the Cartier operator. Let Γ be a curve defined over a perfect field k , let $L = k(\Gamma)$, and let t be an element of L that is transcendental over k and such that $L/k(t)$ is separable. Then any differential ω of L/k is written uniquely as $\omega = (f_0^p + f_1^p t + \dots + f_{p-1}^p t^{p-1}) dt$, where $f_i \in L$. Then the **Cartier operator** C given by $C\omega = f_{p-1} dt$ is well defined and independent of the choice of t and leaves G_1 invariant. Hence given a basis $\omega_1, \dots, \omega_g$ of G_1 , we obtain a matrix (a_{ij}) with coefficient in L by $C\omega_i = \sum a_{ij} \omega_j$ ($1 \leq i \leq g$). This $g \times g$ matrix A is called the **Hasse-Witt matrix** of Γ . The class of A modulo the transformations of the form $S^{-p}AS$ is a birational invariant of Γ and plays an important role in the theory of unramified cyclic p -extensions of the algebraic function field.

F. Generalized Jacobian Varieties

The notion of linear equivalence of divisors on a nonsingular curve can be extended to a more general situation. Such attempts have been made by M. Noether, F. Severi, and M. Rosenlicht, who succeeded in obtaining such a generalization [22].

Let Γ be an algebraic curve, and let P_1, \dots, P_t be points of Γ . Let \mathfrak{D}_{P_i} be the \dagger local ring of P_i . We set $\mathfrak{D} = \bigcap_{i=1}^t \mathfrak{D}_{P_i}$ and $\Gamma' = \Gamma - \{P_1, \dots, P_t\}$. An element of the free Abelian group $G(\Gamma)$ generated by points of Γ' is called a Γ -divisor. Let \mathfrak{a} be a Γ -divisor and set $L(\mathfrak{a}) = \{f \in \mathfrak{D} \mid (f) + \mathfrak{a} \geq 0\} \cup \{0\}$. Then $L(\mathfrak{a})$ is a finite-dimensional linear space (over the universal domain). The dimension of $L(\mathfrak{a})$ is denoted by $l(\mathfrak{a})$, and we set $\dim |a| = \bar{l}(\mathfrak{a}) - 1$. The upper bound π of $\deg(a) = \bar{\dim} |a|$ is a nonnegative integer and is called the **\mathfrak{D} -genus**

of Γ . We call $\bar{l}(\mathfrak{a}) = \pi - \deg \mathfrak{a} + \bar{\dim} |a|$ the **\mathfrak{D} -speciality index** of the divisor \mathfrak{a} . Let C be a nonsingular curve birationally equivalent to Γ , and let Q_1, \dots, Q_s be points of C that correspond to singular points of Γ . An **\mathfrak{D} -differential** ω is a differential form on C (of $\mathbf{K}(\Gamma) = \mathbf{K}(C)$) such that $\sum_{i=1}^s \text{Res}_{Q_i} f\omega = 0$ for any $f \in \mathfrak{D}$. Then $i(\mathfrak{a})$ is equal to the number of linearly independent \mathfrak{D} -differentials ω such that $(\omega) \geq \mathfrak{a}$ in Γ' . The equality $l(\mathfrak{a}) = \deg \mathfrak{a} - \pi + 1 + i(\mathfrak{a})$ is called the **generalized Riemann-Roch theorem**. An \mathfrak{D} -differential ω is called an **\mathfrak{D} -differential of the first kind** if ω is regular everywhere on Γ' . The number of linearly independent \mathfrak{D} -differentials of the first kind is equal to the \mathfrak{D} -genus π . Let g be the effective genus of Γ , i.e., the genus of C . Then we have the equality $\pi - g = \dim_k(\bar{\mathfrak{D}}/\mathfrak{D}) = \delta$, where $\bar{\mathfrak{D}}$ is the integral closure of \mathfrak{D} in $\mathbf{K}(\Gamma)$. The set of \mathfrak{D} -differentials forms an \mathfrak{D} -module that is in general not of rank 1. Hence in this case, we do not have the "canonical divisor." Let c be the conductor of $\mathfrak{D}/\mathfrak{D}$. Then c determines a divisor in a natural way. If we denote the degree of this divisor by d , we have the inequality $\delta + 1 \leq d \leq 2\delta$. We have $d = 2\delta$ if and only if the set of \mathfrak{D} -differentials forms an \mathfrak{D} -module of rank 1. This case occurs, for example, if Γ is a curve on a nonsingular surface or a complete intersection. Two Γ -divisors \mathfrak{a} and \mathfrak{b} are said to be **\mathfrak{D} -linearly equivalent** if there exists a unit f of \mathfrak{D} such that $\mathfrak{a} - \mathfrak{b} = (f)$. The set of Γ -divisors that are \mathfrak{D} -linearly equivalent to zero forms a subgroup $\bar{G}_t(\Gamma)$ of $G(\Gamma)$. There exists a group variety $J_{\mathfrak{D}}$, unique up to isomorphism, that satisfies the four conditions required for Jacobian varieties (Section E) with respect to the class group $\bar{G}_0(\Gamma)/\bar{G}_t(\Gamma)$. The variety J_0 is called the **generalized Jacobian variety**. The generalized Jacobian variety is not complete, in general. If J is the Jacobian variety of C , then $J_{\mathfrak{D}}$ is an extension of J by a connected \dagger linear algebraic group $I_{\mathfrak{D}}$. Any Abelian extension of the function field of Γ can be obtained by the \dagger isogenies of the generalized Jacobian variety of Γ [22]. This fact plays an important role in class field theory over algebraic function fields (\rightarrow 59 Class Field Theory G). The theory for nonsingular curves is considered as the special case in which $\mathfrak{D} = \mathbf{K}(\Gamma)$.

Suppose that Γ is situated in a projective space of dimension n . Let \mathfrak{p} be the prime ideal in $k[X_0, X_1, \dots, X_n]$ defining Γ and $\chi(\mathfrak{p}, m)$ be the number of linearly independent homogeneous polynomials of degree m modulo \mathfrak{p} . Then $\chi(\mathfrak{p}, m)$ is a polynomial in m for large m . This polynomial is called the **Hilbert polynomial** of \mathfrak{p} (or Γ). Let c be the constant term of the Hilbert polynomial. The number $p_a(\Gamma) = 1 - c$ is called the **arithmetic genus** of Γ and is equal to the \mathfrak{D} -genus of Γ . Let Γ be a nonsingular

irreducible curve in \mathbf{P}^3 of degree d . If Γ is contained in a plane, then $p_a(\Gamma) = (d-1) \cdot (d-2)/2$. Otherwise, $p_a(\Gamma) \leq d^2/4 - d + 1$ when d is even, and $p_a(\Gamma) \leq (d^2 - 1)/4 - d + 1$ when d is odd [11].

G. Sheaf Theory

Let Γ be an irreducible curve and \mathcal{O}_p be the local ring of p . Then $\mathcal{O}_\Gamma = \bigcup \mathcal{O}_p$ is an algebraic coherent sheaf, which is called the structure sheaf of Γ , and $\dim H^1(\Gamma, \mathcal{O}_\Gamma)$ is equal to the arithmetic genus π of Γ . Let α be a Γ -divisor, and let $\mathcal{O}_\Gamma(\alpha)$ be the sheaf of germs of rational functions f such that $(f) + \alpha \geq 0$ and $f \in \mathcal{O}_Q$ for every singular point Q of Γ (- 383 Sheaves D). Then $\dim H^1(\Gamma, \mathcal{O}_\Gamma(\alpha))$ is equal to the speciality index $\bar{i}(\alpha)$, and $\dim H^0(\Gamma, \mathcal{O}_\Gamma(\alpha))$ is equal to $I(\alpha)$. When Γ has no singular point, the Riemann-Roch theorem is deduced naturally from Serre's duality theorem: $H^1(\Gamma, \mathcal{O}_\Gamma(\alpha)) \cong H^0(\Gamma, \mathcal{O}_\Gamma(-\alpha))$.

H. Algebraic Correspondence

Let Γ be a nonsingular curve. A divisor of the product variety $\Gamma \times \Gamma$ is called an **algebraic correspondence** of Γ [26, 27]. Let D_0 be the subgroup consisting of divisors that are linearly equivalent to degenerate divisors $a \times \Gamma + \Gamma \times b$. Then the class group $\mathcal{C}(\Gamma) = G(\Gamma \times \Gamma)/D_0$ is called the **group of classes of algebraic correspondences**. We write $X \equiv 0$ if X is an element of D_0 . Let X be an algebraic correspondence, k a field of definition for Γ over which X is rational, and P a generic point of Γ over k . Then $X(P) = \text{pr}_2[X(P \times \Gamma)]$ is rational over $k(P)$. The composite $X_1 \circ X_2$ of two correspondences X_1 and X_2 is defined by $(X_1 \circ X_2)(P) = X_1(X_2(P))$ whenever they have meaning. The composite $X_1 \circ X_2$ determines an element of $\mathcal{C}(\Gamma)$ that depends only on the classes of X_1 and X_2 . This multiplication supplies the group $\mathcal{C}(\Gamma)$ with the structure of an associative ring. This ring is called the **correspondence ring** of Γ . The correspondence ring $\mathcal{C}(\Gamma)$ and the ring \mathcal{A} of endomorphisms of the Jacobian variety J are isomorphic, and the isomorphism is given by the following rule: Let ξ be an element of $\mathcal{C}(\Gamma)$, and let X be a divisor in ξ . Let P be a generic point of Γ with reference to k over which X is rational. Let P_0 be a k -rational point of Γ . Then the class of $X(P) - X(P_0)$ modulo $G_t(\Gamma)$ is independent of the choice of a divisor X in the given class. We set $\Psi(P) = \Phi(X(P) - X(P_0))$ and let λ be the linear extension of Ψ . The correspondence $\xi \rightarrow \lambda$ is an anti-isomorphism of $\mathcal{C}(\Gamma)$ and \mathcal{A} . Now we set $\mathcal{A}_0 = \mathcal{A} \otimes \mathbb{Q}$. Then \mathcal{A}_0 contains an automorphism t of order 2 called an **in-**

volution. Let l be a rational prime different from the characteristic p . Then \mathcal{A} has a faithful representation by $2g \times 2g$ matrices with coefficients in l -adic integers. The trace σ of this representation has the property that $\sigma(\beta \otimes \beta') > 0$ if $\beta \neq 0$ (Castelnuovo's lemma). \mathcal{A}_0 is an algebra of finite rank over \mathbb{Q} , and \mathcal{A} is a finitely generated Abelian group. Based on these results A. Weil proved the Riemann hypothesis for congruent ζ -functions on a nonsingular curve (- 450 Zeta Functions P).

I. Coverings

Let Γ and C be nonsingular curves such that there exists a regular mapping $\pi: \Gamma \rightarrow C$. Then there is an injection of the function field $\mathbf{K}(C)$ into $\mathbf{K}(\Gamma)$. If $\mathbf{K}(\Gamma)$ is separably algebraic over $\mathbf{K}(C)$, then Γ is called a **covering (curve)** of C . The integer $[\mathbf{K}(\Gamma): \mathbf{K}(C)] = n$ is called the **degree of covering**. Let P be a point of Γ and let $Q = \pi(P)$. Let t, s be local parameters at P on Γ and at Q on C , respectively. The nonnegative integer $v_P(ds/dt)$ is called the **differential index** at P and is denoted by m_P . The index m_P is zero except for a finite number of points. The divisor $\sum m_P P$ is called the **branch divisor**. The covering Γ is called an **unramified covering** if the branch divisor is zero. If we denote the branch divisor by α , we have the formula $2g(\Gamma) - 2 = n(2g(C) - 2) + \deg \alpha$, where $g(\Gamma)$ and $g(C)$ are genera of Γ and C , respectively. This is called the **Riemann-Hurwitz formula**. This formula yields at once that a rational curve has no nontrivial unramified covering and that Γ can be an unramified covering of itself if and only if Γ is an elliptic curve.

J. Theory of Moduli

Let $\mathcal{M}_g(k)$ be the set of isomorphism classes of complete nonsingular irreducible curves (here simply called "curves") of genus g defined over a field k . We can endow $\mathcal{M}_g(k)$ with a structure of an algebraic variety over k with the property that for any smooth family (over k) $\pi: \mathcal{C} \rightarrow S$ of curves of genus y the map $T: S \rightarrow \mathcal{M}_g(k)$ sending s to the isomorphism class of the fiber $\pi^{-1}(s)$ is a morphism. This algebraic variety is called the **(coarse) moduli space of curves of genus g** . Furthermore, the moduli space over $\text{Spec}(\mathbb{Z})$ exists (D. Mumford [16]). $\mathcal{M}_g(k)$ is normal, irreducible and quasiprojective (but not complete for $g > 0$) of dimension 0 ($g = 0$), 1 ($g = 1$), $3g - 3$ ($g \geq 2$) ($= 3g - 3 + \dim \text{Aut}(C)$) with only quotient singularity [5].

Since when $g = 1$ any elliptic curve C over an algebraically closed field k is isomorphic to a plane curve with the Weierstrass canonical

form $y^2 = 4x^3 - g_2x - g_3$ (if $\text{char}(k) \neq 2$; the case of $\text{char}(k) = 2$ needs a slight modification), the correspondence $C \rightarrow j(g_2, g_3)$ (\rightarrow Section D) defines an isomorphism $\mathcal{M}_1 \simeq \mathbf{A}_k^1$ (affine line). In the case of $g \geq 2$ and $k = \mathbf{C}$ we have another function-theoretic construction of \mathcal{M}_g due to Teichmüller (\rightarrow 11 Algebraic Functions F).

Let \mathcal{A}_g be the coarse moduli space (over \mathbf{Z}) of principally polarized Abelian varieties of dimension g (\rightarrow 3 Abelian Varieties 1). For a curve C we denote the Jacobian variety of C by $J(C)$ (\rightarrow Section E). The correspondence $C \rightarrow J(C)$ defines a morphism $i: \mathcal{M}_g \rightarrow \mathcal{A}_g$ which is injective (Theorem of Torelli). It is even an immersion (F. Oort, J. Steenbrink). If $g = 1, 2, 3$, then i is an open immersion whose image we can describe.

For $k = \mathbf{C}$ we can express the above map i by using periods of curves. Namely, let $\alpha_1, \dots, \alpha_g, \beta_1, \dots, \beta_g$ be a (canonical) basis of $H_1(C, \mathbf{Z})$ defined by normal sections of C (considered as a real surface). Let $\omega_1, \dots, \omega_g$ be a basis of differential forms of the first kind on C with $\int_{\beta_i} \omega_j = \delta_{ij}$. Then the matrix $\Omega = (\int_{\alpha_i} \omega_j)$ is symmetric and has a positive definite imaginary part, i.e., is an element of the Siegel upper half-space \mathfrak{S}_g of degree y . With the identification $\mathcal{A}_g \simeq \mathfrak{S}_g / \text{Sp}(g, \mathbf{Z})$ (\rightarrow 3 Abelian Varieties 1) the map $i: \mathcal{M}_g \rightarrow \mathcal{A}_g$ is nothing but the one sending the isomorphism class of C to $\text{Rmod } \text{Sp}(g, \mathbf{Z})$. For $g = 4$ the closure of $i(\mathcal{M}_4)$ is a principal divisor defined explicitly with theta constants (Schottky, J.-I. Igusa).

K. Stable Curves

A reduced connected complete curve C over k is called a **stable curve** of genus g (> 0) if (i) C has only ordinary double points as possible singularities; (ii) when Γ is a smooth rational irreducible component of C , then Γ intersects the other components in more than 2 points; (iii) $\dim, H^1(C, \mathcal{O}_C) = g$.

There exists a coarse moduli space \mathcal{S}_g of stable curves of genus g which contains \mathcal{M}_g as a Zariski open subset. The space \mathcal{S}_g defined over $\text{Spec}(\mathbf{Z})$ is complete and even projective; hence \mathcal{S}_g gives a compactification of \mathcal{M}_g . The completeness of \mathcal{S}_g follows from the **stable reduction theorem** as follows: let R be a discrete valuation ring with quotient field K and C be a smooth connected curve over K . Then there exists a finite separable algebraic extension L of K such that the curve $C \times_K L$ extends to a flat family of stable curves over $\text{Spec}(R_L)$, where R_L denotes the integral closure of R in L . In this case we say that the curve $C \times_K L$ has a **stable reduction** in R . With the above notation a curve C over K has a stable reduction in R if and only if its

Jacobian variety $J(C)$ has a stable reduction in R (\rightarrow 3 Abelian Varieties N).

Over \mathbf{C} , \mathcal{A}_g has the Satake compactification \mathcal{A}_g , which is a disjoint union of $\mathcal{A}_{g'}, 0 \leq g' \leq g$, as a set. The injection $i: \mathcal{M}_g \rightarrow \mathcal{A}_g$ (\rightarrow Section J) extends to a morphism $j: \mathcal{S}_g \rightarrow \mathcal{A}_g$ that sends the isomorphism class of a stable curve $C = C_1 \cup \dots \cup C_k$ (C_i irreducible) to that of $J(\tilde{C}_1) \times \dots \times J(\tilde{C}_k) \in \mathcal{A}_{g'}$, $g' = \sum_i \text{genus}(\tilde{C}_i)$, where \tilde{C}_i is the normalization of C_i . In particular the closure of the image of i in \mathcal{A}_g is the set of products of Jacobian varieties.

References

- [1] S. Arakelov, Families of algebraic curves with fixed degeneracies, Math. USSR-I-V., 5 (1971), 1277-1302.
- [2] H. F. Baker, Analytical principles of the theory of curves, Principles of Geometry 5, Cambridge Univ. Press, 1933.
- [3] A. Brill and M. Noether, Über die algebraischen Functionen und ihre Anwendungen in der Geometrie, Math. Ann., 7 (1874), 269-310.
- [4] G. Castelnuovo, Numero delle involuizioni razionali giacenti sopra una curva di dato genere, Rend. Acad. Lincei, 4, 5 (1889).
- [5] P. Deligne and D. Mumford, The irreducibility of the space of curves of given genus, Publ. Math. Inst. HES, 36 (1969), 75-109.
- [6] F. Enriques and O. Chisini, Teoria geometrica delle equazioni delle funzioni algebriche, Zanichelli, Bologna, 1934.
- [7] T. Fujita, Defining equations for certain types of polarized varieties, Complex Analysis and Algebraic Geometry, Iwanami, Tokyo, 1977, 165-173.
- [8] W. Fulton, Algebraic curves, Benjamin, 1969.
- [9] P. Griffiths and J. Harris, Principles of algebraic geometry, Wiley, 1978.
- [10] P. Griffiths and J. Harris, On the variety of special linear systems on a general algebraic curve, Duke Math. J., 47 (1980), 233-272.
- [11] R. Hartshorne, Algebraic geometry, Graduate text 52, Springer, 1977.
- [12] R. Hartshorne, On the classification of algebraic space curves, Vector Bundles and Differential Equations, Birkhauser, 1980.
- [13] G. Kempf, Schubert methods with an application to algebraic curves, Publications of Mathematisch Centrum, Amsterdam, 1971.
- [14] S. Kleiman and D. Laksov, On the existence of special divisors, Amer. J. Math., 94 (1972), 431-436.
- [15] S. Kleiman, r-Special subschemes and an argument of Severi's, Advances in Math., 22 (1976), 1-31.

- [16] D. Mumford, Geometric invariant theory, Erg. Math., Springer, 1965.
- [17] D. Mumford, Curves and their Jacobians, Univ. of Michigan Press, 1975.
- [18] M. Namba, Families of meromorphic functions on compact Riemann surfaces, Lecture notes in math. 767, Springer, 1979.
- [19] M. Noether, Zur Grundlegung der Theorie der algebraischen Raumcurven, J. Reine Angew. Math., 93 (1882), 271-318 (Berlin, 1883).
- [20] B. Saint-Donat, On Petri's analysis of the linear system of quadrics through a canonical curve, Math. Ann., 206 (1973), 157-175.
- [21] A. Seidenberg, Elements of the theory of algebraic curves, Addison-Wesley, 1968.
- [22] J.-P. Serre, Groupes algébriques et corps de classes, Actualités Sci. Ind., Hermann, 1959.
- [23] F. Severi, Vorlesungen über algebraische Geometrie, Teubner, 1921.
- [24] J. M. Wahl, Deformations of plane curves with nodes and cusps, Am. J. Math., 96 (4) (1974), 529-577.
- [25] R. J. Walker, Algebraic curves, Princeton Univ. Press, 1950.
- [26] A. Weil, Sur les courbes algébriques et les variétés qui s'en déduisent, Actualités Sci. Ind., Hermann, 1948.
- [27] A. Weil, Variétés abéliennes et courbes algébriques, Actualités Sci. Ind., Hermann, 1948.
- [28] A. Weil, Zum Beweis des Torellischen Satzes, Nachr. Akad. Wiss. Göttingen (1957), 33-53.
- [29] H. Weyl, Die Idee der Riemannschen Fläche, Teubner, third edition, 1955.
- [30] G. B. Winters, On the existence of certain families of curves, Amer. J. Math., 96 (1974), 215-228.

10 (111.5) Algebraic Equations

A. General Remarks

Let $F_1(X_1, \dots, X_r), \dots, F_r(X_1, \dots, X_r)$ be r tpolynomial in m variables X_1, \dots, X_m over a \dagger field k . Then the equations

$$F_1 = 0, \dots, F_r = 0$$

are called **algebraic equations in m unknowns**. When we consider these equations simultaneously, where $r \geq 2$, we call them a **system of r equations** or **simultaneous equations**. (For $r = 1$, a system of one equation means the single equation $F_1 = 0$.) Coefficients of F_1, \dots, F_r are called **coefficients** of the system, and the great-

est of the degrees of F_1, \dots, F_r is called the **degree** of the system.

To **solve** a system of equations (henceforth in this article we shall omit the word "algebraic") means to find the common \dagger zero points (in an \dagger algebraically closed field containing k) of elements F_1, \dots, F_r of the \dagger polynomial ring $k[X_1, \dots, X_r]$. If there exist no common zero points, the system is said to be **inconsistent**; if there exists a finite number of such points, it is said to be **regular**; and if there are an infinite number of such points, it is called **indeterminate**. The elimination method allows us to reduce the problem of solving a system of r equations to the case $r = 1$. In particular, any regular system of equations can be reduced to the case $m=r=1$.

B. Equations in One Unknown

For the above reason, it is important to consider an equation of the form $f(X) = 0$, where

$$f(X) = a_0 X^n + a_1 X^{n-1} + \dots + a_n, \quad a_0 \neq 0. \quad (1)$$

This gives the general form of an algebraic equation in one unknown.

According as $f(X)$ is reducible or not in the \dagger polynomial ring $k[X]$, the equation $f(X)=0$ is called **reducible** or **irreducible** (- 337 Polynomials). In some \dagger algebraic extension field K of k , $f(X)$ can be factored:

$$f(X) = a_0(X - \alpha_1)(X - \alpha_2) \dots (X - \alpha_n). \quad (2)$$

$\alpha_1, \dots, \alpha_n$ are called the **roots** of the equation $f(X) = 0$. Hence, any algebraic equation of degree n has exactly n roots (**Kronecker's theorem**). Now, $(-1)^i a_i / a_0$ is equal to the \dagger elementary symmetric function of degree i of $\alpha_1, \dots, \alpha_n$. Some of the roots $\alpha_1, \dots, \alpha_n$ may be identical. If α appears ρ times in $\alpha_1, \dots, \alpha_n$, we say that α is a ρ -tuple root, and ρ is called the **multiplicity** of the root α . When $\rho = 1$, α is called a **simple root**, and when $\rho \geq 2$, α is called a **multiple root**. Let β_1, \dots, β_v be all the distinct roots among $\alpha_1, \dots, \alpha_n$, and let ρ_i be the multiplicity of β_i ($i = 1, \dots, v$). Then

$$f(X) = a_0(X - \beta_1)^{\rho_1} \dots (X - \beta_v)^{\rho_v}, \quad (2')$$

$$\rho_1 + \dots + \rho_v = n.$$

If ρ_1, \dots, ρ_v are not divisible by the \dagger characteristic of k , the greatest common divisor g of f and

$$f' = na_0 X^{n-1} + (n-1)a_1 X^{n-2} + \dots + a_{n-1}$$

is $(X - \beta_1)^{\rho_1-1} \dots (X - \beta_v)^{\rho_v-1}$. Thus we can reduce the multiplicity of every root to 1 by dividing f by g . Any irreducible equation over a field of characteristic 0 has no multiple roots. Equation (1) has multiple roots if and only if

its [†]discriminant D is equal to 0 (- 149 Fields; 172 Galois Theory).

C. Equations of Special Type

In Sections C and D we assume that the characteristic of k is zero.

Binomial Equations. An equation of the type $X^m - a = 0$ is called a **binomial equation**. It is solved by **root extraction**. Let $\sqrt[m]{a}$ (mth root of a) be one of the roots (if a is a positive real number, $\sqrt[m]{a}$ usually denotes a positive real root). Then $\sqrt[m]{a}$ multiplied by $1, \zeta, \zeta^2, \dots, \zeta^{m-1}$ are the roots of $X^m - a = 0$, where ζ is a [†]primitive mth root of unity.

Reciprocal Equations. An equation $a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0 = 0$ is called a **reciprocal equation** if $a_0 = a_n, a_1 = a_{n-1}, a_2 = a_{n-2}, \dots$. A reciprocal equation of an odd degree $n = 2m + 1$ has a root $X = -1$, and dividing the left side by $X + 1$ we get a reciprocal equation of degree $2m$. A reciprocal equation of degree $n = 2m$ is reduced to an equation of degree m in $Y = X + X^{-1}$ and the quadratic equation $X^2 - XY + 1 = 0$.

D. Equations of Lower Degree

(- Appendix A, Table 1)

(1) A **linear equation** $a_0 X + a_1 = 0$ has a single root $-a_1/a_0$. (2) the roots of a **quadratic equation** $a_0 X^2 + a_1 X + a_2 = 0$ are given by $(-a_1 \pm \sqrt{a_1^2 - 4a_0 a_2})/2a_0$. (3) To solve a **cubic equation** $a_0 X^3 + a_1 X^2 + a_2 X + a_3 = 0$, we set $A_1 = 9a_0 a_1 a_2 - 2a_1^3 - 27a_0^2 a_3, A_2 = a_1^2 - 3a_0 a_2$, and solve the quadratic equation $T^2 - A_1 T + A_2 = 0$. Let t_1 and t_2 be the roots of this quadratic equation, and let ω be any cube root of 1. Then $(-a_1 + \omega \sqrt[3]{t_1} + \omega^2 \sqrt[3]{t_2})/3a_0$ is a root of the original cubic equation. (**Cardano's formula**). If we apply this method to a cubic equation $aX^3 + bX^2 + cX + d = 0$ with real coefficients, we need to use complex cube roots even if the roots of the equation are real. In fact, it has been proved that it is not possible to solve this equation within the real numbers in this case; i.e., if the cubic equation is irreducible over the extension $\mathbf{Q}(a, b, c, d)$ of the rational number field \mathbf{Q} , and if all of its roots are real, it is impossible to find the roots only by rational operations and with real radicals. This is called the **casus irreducibilis**. (4) A **quartic equation** $a_0 X^4 + a_1 X^3 + a_2 X^2 + a_3 X + a_4 = 0$ can be solved by means of reduction to a cubic equation (L. Ferrari) (- Appendix A, Table 1). Generally, the procedure of solving an algebraic equation, i.e.,

finding the roots of a given equation from its coefficients by means of a finite number of rational operations and extractions of radicals, is called a **solution by radicals** (or **algebraic solution**). The [†]general algebraic equation whose degree is ≥ 5 cannot be solved by radicals (N. H. Abel) (- 172 Galois Theory).

E. Analytic Theory

In this section, k denotes the field \mathbf{R} of real numbers or the field \mathbf{C} of complex numbers. These cases have been studied for a long time, for practical reasons.

Concerning the case $k = \mathbf{C}$, the field \mathbf{C} is [†]algebraically closed; i.e., every equation with coefficients from \mathbf{C} has a root in \mathbf{C} (**Gauss's theorem**, called the **fundamental theorem of algebra**). Accordingly, in the field \mathbf{C} , we always have equations (2) and (2').

Let $\alpha_1, \dots, \alpha_n$ be the roots of equation (1). Then each α_i is a continuous function of coefficients a, a_1, \dots, a_n . Concerning the location of roots of $f(X) = 0$ and $f'(X) = 0$ on the complex plane, we have the following theorems:

(1) Any convex polygon on the complex plane containing the roots of $f(X) = 0$ also contains the roots of $f'(X) = 0$ (Gauss).

(2) Let C be a rectifiable [†]Jordan curve not passing through a root of $f(X) = 0$. Then the number (C, f) of the roots of $f(X) = 0$ lying in the region enclosed by C is equal to $(1/2\pi i) \int_C (f'(z)/f(z)) dz$, where the multiplicity of the roots is taken into account.

(3) Let C be a Jordan curve on the complex plane. If $|f(z)| > |g(z)|$ at every point z on C , then the equations $f = 0$ and $f + g = 0$ have the same number of roots (counting multiplicity) within the region enclosed by C (**Rouché's theorem**).

(4) The absolute value of a root of equation (1) is less than

$$M = \max(|a_1/a_0|, \dots, |a_n/a_0|) + 1.$$

(5) Let D be the [†]discriminant of f , and assume that $|\alpha_i| \leq M$ ($i = 1, \dots, n$). Then $|\alpha_i - \alpha_j|^2 \geq D/(2M)^{n(n-1)-2} = E$. Since the value of $|D|$ is known from [†]and one value of M is given by theorem (4), we have one value of E . If we draw a circle on the complex plane with center at the origin and with radius M , and if we cover it with a net whose meshes have diameters less than $\sqrt{E}/2$, then the interior of each mesh contains at most one root of $f = 0$.

When $k = \mathbf{R}$, i.e., $f \in \mathbf{R}[X]$, let β_1, \dots, β_v denote the distinct roots of $f = 0$, and recall equation (2'). Suppose that $\beta_1, \dots, \beta_\lambda \in \mathbf{R}$ and the others $\notin \mathbf{R}$. Then $v - \lambda$ is an even integer 2κ , and we can renumber $\beta_{\lambda+1}, \dots, \beta_v$ so that $\beta_{\lambda+1} = \beta_{\lambda+\kappa+1}, \dots, \beta_{\lambda+\kappa} = \beta_v$ (β denotes the con-

jugate of β) and $\rho_{\lambda+1} = \rho_{\lambda+\kappa+1}, \dots, \rho_{\lambda+\kappa} = \rho_v$. In this case, $\beta_1, \dots, \beta_\lambda$ are the **real roots** of equation (2), and the other β 's are called the **imaginary roots**.

(6) If $f \in \mathbf{R}[X]$ and $a_0 > a_1 > \dots > a_n > 0$, then the absolute value of any root of equation (1) is less than 1 (**Kakeya-Eneström theorem**).

Concerning the real roots of an equation $f = 0$, where $f \in \mathbf{R}[X]$, we have the following theorems: Let $N(a, b)$ ($a, b \in \mathbf{R}$) denote the number of real roots in the interval (a, b) . Furthermore, let $V(c_1, c_2, \dots, c_p)$ denote the number of changes of sign in the sequence c_1, c_2, \dots, c_p of real numbers, which is defined as follows: Suppose that we have the sequence c_{v_1}, \dots, c_{v_q} after deleting the terms $c_i = 0$ from the sequence c_1, c_2, \dots, c_p . Then

$$V(c_1, c_2, \dots, c_p) = \frac{1}{2} \sum_{j=1}^{q-1} (1 - \operatorname{sgn} c_{v_j} c_{v_{j+1}}).$$

(7) $N(0, \infty) \equiv V(a_0, a_1, \dots, a_n) \pmod{2}$ and $N \leq V$ (**Descartes's theorem**).

(8) Let $V(c) = V(f(c), f'(c), \dots, f^{(n)}(c))$. Then $N(a, b) \equiv V(a) - V(b) \pmod{2}$ and $N \leq V(a) - V(b)$ (**Fourier's theorem**).

(9) We may assume that $f = 0$ has no multiple roots. Construct a finite series $f_0 = f, f_1 = f', \dots, f_l$ of polynomials over \mathbf{R} such that $f_{i-1} = f_i q_i - f_{i+1}$ for $i = 1, 2, \dots, l-1$ and $f_l \in \mathbf{R}$, by successive application of the \dagger division algorithm. Let $V(c) = V(f_0(c), f_1(c), \dots, f_l(c))$. Then $N(a, b) = V(a) - V(b)$ (**Sturm's theorem**). By means of this theorem we can determine the location of real roots as precisely as we wish.

(10) In order that every root α_i of an equation $f = 0$ with $a_n > 0$ lies on the left side of the imaginary axis, i.e., $\operatorname{Re} \alpha_i < 0$, it is necessary and sufficient that in the following matrix the \dagger principal minors composed of the first r rows and first r columns be positive for all $r = 1, 2, \dots, n$ (**Hurwitz's theorem**):

$$\begin{pmatrix} a_n & a_{n-1} & a_{n-2} & a_{n-3} & a_{n-4} & \dots \\ 0 & a_n & a_{n-1} & a_{n-2} & a_{n-3} & \dots \\ 0 & 0 & a_n & a_{n-1} & a_{n-2} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_{n-2} & a_{n-3} & \dots \end{pmatrix}.$$

Also, for $f \in \mathbf{C}[X]$, various results have been obtained about under what conditions all the roots of $f = 0$ lie on one side of a given straight line or inside a given circle (e.g., the unit circle) (- 301 Numerical Solution of Algebraic Equations).

References

- [1] L. E. Dickson, Elementary theory of equations, Wiley, 1914.
- [2] J. Dieudonné, La théorie analytique des

polynômes, Mémor. Sci. Math., Gauthier-Villars, 1938.

[3] M. Marden, Geometry of polynomials, Amer. Math. Soc. Math. Surveys, second edition, 1966.

[4] W. Specht, Algebraische Gleichungen mit reellen oder komplexen Koeffizienten, Teubner, 1958.

11 (X1.10) Algebraic Functions

A. Definition

An **algebraic function** is a multiple-valued \dagger analytic function $w = w(z)$ defined by an irreducible algebraic equation $P(z, w) = 0$ with complex coefficients.

B. History and Methods

The theory of algebraic functions evolved from the works of C. F. Gauss, N. H. Abel, and C. G. J. Jacobi on \dagger elliptic functions in the early 19th Century. Stimulated by their works, B. Riemann and K. Weierstrass established the foundations of the theory of complex functions and developed the important theory of algebraic functions.

The equation $P(z, w) = 0$ defines a curve in the 2-dimensional complex \dagger projective space with inhomogeneous coordinates z, w . Investigations from this point of view were initiated by Riemann, A. Clebsch, and P. Gordan. This approach was followed by A. Brill, M. Noether, and the Italian school (F. Severi, C. Segre, etc.) and has developed into contemporary algebraic geometry (\rightarrow 9 Algebraic Curves, 12 Algebraic Geometry).

The set of \dagger function elements $w(z)$ satisfying $P(z, w) = 0$ is a \dagger complex manifold \mathfrak{R} , a closed (= compact) \dagger Riemann surface, on which z and w are meromorphic functions. The field $K_{\mathfrak{R}}$ consisting of the meromorphic functions on \mathfrak{R} is an algebraic function field $\mathbf{C}(z, w)$. Conversely, for any closed Riemann surface \mathfrak{R} , the field $K_{\mathfrak{R}}$ is an algebraic function field in one variable over \mathbf{C} , and any pair of functions z and w with $K_{\mathfrak{R}} = \mathbf{C}(z, w)$ has the property that \mathfrak{R} is \dagger conformally equivalent to the Riemann surface determined in the above fashion by the irreducible algebraic equation $P(z, w) = 0$ satisfied by z and w . Two Riemann surfaces $\mathfrak{R}_1, \mathfrak{R}_2$ determined by the equations $P_1 = 0, P_2 = 0$ are conformally equivalent if and only if the fields $K_{\mathfrak{R}_1}$ and $K_{\mathfrak{R}_2}$ are \mathbf{C} -isomorphic. This condition is equivalent to the existence of a birational

transformation between the algebraic curves $P_1 = 0, P_2 = 0$. The "analytic method" (the method of studying algebraic functions as functions on Riemann surfaces) is the creation of Riemann. It was extended by F. Klein and D. Hilbert, and later by H. Weyl, who established in his monograph [6] a rigorous foundation of the analytic method for the theory of algebraic functions.

Given an arbitrary algebraic function field K in one variable over C , the set \mathfrak{R} of its \dagger prime divisors with a suitable topology and analytic structure is a closed Riemann surface whose function field $K_{\mathfrak{R}}$ coincides with K . The "algebraic method" (the method of studying algebraic functions as elements of an algebraic function field) was founded by J. W. Dedekind and H. Weber at the end of the 19th Century. In the 20th Century, the algebraic method has made remarkable progress, owing to the development of abstract algebra. It covers the case of an arbitrary ground field as well as that of more than one variable. The theory of algebraic functions has had considerable influence on the development of number theory because of a basic analogy between the two subjects.

The \dagger universal covering spaces (surface) $\tilde{\mathfrak{R}}$ of a closed Riemann surface \mathfrak{R} can be regarded, by conformal mapping, as the Riemann sphere, the plane, or the unit disk (or, equivalently, to the upper half-plane) if the \dagger genus g of \mathfrak{R} is 0, 1, or ≥ 2 , respectively. Then the \dagger covering transformation group G , consisting of \dagger linear fractional transformations without fixed points in $\tilde{\mathfrak{R}}$, is \dagger properly discontinuous and has a compact \dagger fundamental domain. Conversely, if D is one of the three domains just mentioned and if G is the group just described, then $\mathfrak{R} = D/G$ is a closed Riemann surface such that D and G are its universal covering space and covering transformation group. A meromorphic function on \mathfrak{R} is represented as an automorphic function on $\tilde{\mathfrak{R}}$ with respect to G . If $y = 0$, then $G = \{1\}$, $\tilde{\mathfrak{R}} = \mathfrak{R}$, and $K_{\mathfrak{R}}$ is the field of rational functions. If $g = 1$, then $K_{\mathfrak{R}}$ is the field of \dagger elliptic functions. The study of algebraic functions as automorphic functions was initiated by H. Poincaré and Klein. Recently, C. L. Siegel made a remarkable contribution to the investigation of the case of several variables. The theory of automorphic functions is also related to number theory. Works of E. Hecke, M. Eichler, and G. Shimura on this domain are noteworthy (- 32 Automorphic Functions, 73 Complex Multiplication).

Another important topic concerning algebraic functions (closed Riemann surfaces) is the problem of moduli. Riemann stated, without rigorous proof, that the set of conformal

equivalence classes of closed Riemann surfaces of genus $g (\geq 2)$ depends on $3g - 3$ complex parameters, called **moduli**. This has led to the modern theory of Teichmüller spaces, which is developing into an extensive new field (- 234 Kleinian Groups, 416 Teichmüller Spaces).

In the rest of this article, we deal mainly with the analytic method (for the case of two variables $\rightarrow [1]$).

C. Abelian Differentials

An **Abelian differential** on a closed Riemann surface \mathfrak{R} is, by definition, a complex \dagger differential form $\omega = a(z)dz$, where $a(z)$ is a meromorphic function of a local parameter z . Such a differential is said to be of the **first kind** if $a(z)$ is holomorphic, of the **second kind** if the residue vanishes everywhere, and of the **third kind** otherwise.

The indefinite integral $W(p) = \int_{p_0}^p \omega$ of an Abelian differential ω , where p_0 is assumed not to be a pole of ω , is called an **Abelian integral**. It is said to be of the **first, second, or third kind** if the same holds for ω . If y is a \dagger 1-cycle on \mathfrak{R} , the quantity $\int_y \omega$ is referred to as the **period** of ω along y . An **elliptic integral** is defined to be an Abelian integral on a closed Riemann surface of genus 1. For example, this is the case if the equation $P(z, w) = 0$ defining the surface is of degree 2 with respect to w and of degree 3 or 4 with respect to z . More generally, a closed Riemann surface is called **hyperelliptic** if $P(z, w)$ is of degree 2 with respect to w , or, equivalently, if \mathfrak{R} carries a meromorphic function with exactly two poles. An Abelian integral on such a surface is called a **hyperelliptic integral**.

On a closed surface \mathfrak{R} , let V_a be the linear space over C of the Abelian differentials of the first kind. Given a 1-cycle α of \mathfrak{R} , there exists a unique $\omega_\alpha \in V_a$ such that $\text{Re} \int_y \omega_\alpha$ is equal to the intersection number (α, y) for every 1-cycle y . This differential is also characterized by the property $((\omega, \omega_\alpha)_\mathfrak{R} \equiv) \int_\mathfrak{R} \omega \wedge * \bar{\omega}_\alpha = -2\sqrt{-1} \int_\alpha \omega$ for every $\omega \in V_a$. If $\{\alpha_1, \dots, \alpha_{2g}\}$ form a basis of the 1-dimensional homology group with integral coefficients, then $\text{Re} \omega_{\alpha_i} (i = 1, \dots, 2g)$ form a basis of the linear space V_h over R of the \dagger harmonic differentials on \mathfrak{R} as well as that of the space $\{\text{Re} \omega \mid \omega \in V_a\}$. Accordingly,

$$\dim_C V_a = g, \quad \dim V_h = 2g.$$

These identities show a close relationship between the topological structure of \mathfrak{R} and the space of the Abelian differentials on \mathfrak{R} (- 194 Harmonic Integrals).

One can choose a 1-dimensional homology basis $\{\alpha_i, \alpha_{g+i}\}_{i=1}^g$ so that $(\alpha_i, \alpha_j) = (\alpha_{g+i}, \alpha_{g+j}) = 0$, $(\alpha_i, \alpha_{g+i}) = 1$, and $(\alpha_i, \alpha_{g+j}) = 0 (i \neq j), (i, j = 1, 2, \dots, g)$. Such a basis is called a **canonical**

homology basis. If ω and σ are Abelian differentials of the first kind, then we have

$$(\omega, \bar{\sigma})_{\mathfrak{R}} = -\sqrt{-1} \left(\int_{\mathfrak{A}} \omega \int_{\mathfrak{A}_{g+j}} \bar{\sigma} - \int_{\mathfrak{A}_{g+j}} \omega \int_{\mathfrak{A}_j} \bar{\sigma} \right) = 0,$$

$$(\omega, \omega)_{\mathfrak{R}} = \sqrt{-1} \sum_{j=1}^g \left(\int_{\mathfrak{A}_j} \omega \int_{\mathfrak{A}_{g+j}} \bar{\omega} - \int_{\mathfrak{A}_{g+j}} \omega \int_{\mathfrak{A}_j} \bar{\omega} \right) \geq 0.$$

Together, these are called the (Riemann) **period relation**. The second formula implies that ω vanishes identically if the periods of ω along all α_j ($j = 1, \dots, g$) are zero.

Let $\alpha_1, \dots, \alpha_{2g}$ be a 1-dimensional homology basis, and let $\omega_1, \dots, \omega_g$ form a basis of V_a over \mathbb{C} . The $g \times 2g$ matrix Ω with $\int_{\alpha_j} \omega_i$ as its (i, j) -component is called a **period matrix**. Corresponding to the change of bases (a) and (w), it is subject to transformation into the form $A\Omega M$, where A is a $g \times g$ invertible complex matrix and M is a $2g \times 2g$ integral square matrix with determinant ± 1 . Conversely, two Riemann surfaces are conformally equivalent if they possess period matrices transformable to each other in this manner (**Torelli's theorem**). We can choose $\omega_1, \dots, \omega_g$ so that the corresponding period matrix with respect to a canonical homology basis becomes (I_g, T) with the $g \times g$ unit matrix I_g . Then from the period relation, T is symmetric and $\text{Im } T$ is positive definite.

On the complex linear space \mathbb{C}^g , consider the subgroup generated by the $2g$ column vectors of a period matrix Ω (the subgroup is also denoted by Ω). Since it is of rank $2g$ and properly discontinuous, a group manifold \mathbb{C}^g/Ω is obtained. It is determined by \mathfrak{R} uniquely up to analytic isomorphism and is called the **Jacobian variety** of \mathfrak{R} . The **generalized Jacobian variety** is introduced in a similar fashion by means of Abelian integrals of the second and third kinds (- 9 Algebraic Curves).

D. The Riemann-Roch Theorem

In the present context, a 0-chain with integral coefficients on a Riemann surface \mathfrak{R} is referred to as a **divisor**. A divisor $d = \sum n_i p_i$ ($n_i \in \mathbb{Z}$, $p_i \in \mathfrak{R}$) is an **integral divisor** (or **positive divisor**) if $n_i \geq 0$ in the reduced expression; d is a **prime divisor** if it consists of a single point p_1 and $n_1 = 1$. A divisor of a meromorphic function for an Abelian differential ω is defined by taking the p_i as the zeros (poles) off or ω and n_i ($= n_i$) as the multiplicity of the zero

(pole) at p_i . The divisors on \mathfrak{R} constitute an Abelian group \mathfrak{D} in which **principal divisors**, i.e., divisors of meromorphic functions, constitute a subgroup \mathfrak{P} . The factor group $\mathfrak{D}/\mathfrak{P}$ is called the **divisor class group**; an element of it is called a **divisor class**. The divisors of Abelian differential constitute a single divisor class, which is referred to as the **canonical divisor class** (or **differential divisor class**). The **degree** and the **dimension** of a divisor class D are defined as follows, independent of the choice of the representative $d = \sum n_i p_i \in D$: $\deg D = \sum n_i$, $\dim D = \dim_{\mathbb{C}} \{f | f \text{ is meromorphic, (divisor of } f) + d \text{ is a positive divisor}\}$. For example, the degree of the principal divisor class is zero.

In terms of these concepts, the **Riemann-Roch theorem** is stated as follows: For a divisor class D on a closed Riemann surface \mathfrak{R} of genus g and for an integer n , we have

$$\dim(D + nW) - \dim(-D - (1 - n)W) = \deg D + (2n - 1)(g - 1),$$

where W is the canonical divisor class (- 9 Algebraic Curves).

This theorem implies the following properties of \mathfrak{R} : (i) $\deg W = 2g - 2$. (ii) The holomorphic invariant forms φdz^2 (i.e., analytic tensors of order 2), referred to as **quadratic differentials**, constitute a linear space over \mathbb{C} of dimension 0 (if $g = 0$), 1 (if $g = 1$), or of dimension $3g - 3$ (if $g \geq 2$). The quadratic differentials have close connection with extremal quasiconformal mappings and play an important role in the theory of Teichmüller spaces (- 352 Quasiconformal Mappings, 416 Teichmüller Spaces). (iii) For a point $p \in \mathfrak{R}$, a positive integer m is called a **gap value** if \mathfrak{R} carries no meromorphic function having a pole only at p with multiplicity m . Then if $g = 0$, no point has gap values; and if $g \geq 1$, every point p has exactly g gap values; in this case, p also has a nongap value $m \leq g + 1$. A point p is called an **ordinary point** if the gap values at p are $1, 2, \dots, g$; otherwise p is called a **Weierstrass point**. If $g \geq 2$, then the total number N of Weierstrass points is not less than $2g + 2$ and not greater than $(g - 1)g(g + 1)$ (A. Hurwitz). Moreover, the case $N = 2g + 2$ occurs if and only if \mathfrak{R} is hyperelliptic, and then the gap values at Weierstrass points are $1, 3, \dots, 2g - 1$. This implies that every closed Riemann surface of genus 2 is hyperelliptic. (iv) Suppose the genus g of \mathfrak{R} is ≥ 2 . A conformal mapping f of \mathfrak{R} onto itself with the property that every 1-cycle γ is always homologous to $f(\gamma)$ is necessarily the identity transformation. Also, \mathfrak{R} is known to admit only a finite number of conformal mappings onto itself (H. Schwarz); the total number does not exceed $84(g - 1)$ (Hurwitz).

E. Abel's Theorem

Abel's theorem is stated as follows: A divisor d of degree zero is a principal divisor if and only if it is expressed as $d = \partial\gamma$ by means of a 1-chain γ that has the property that $\int_{\gamma} \omega = 0$ for every $\omega \in V_a$.

Given a divisor class D of degree zero, consider a 1-chain γ with $\partial\gamma \in D$. For every 1-cycle α there corresponds the quantity

$$\chi_{\alpha}(D) = \exp\left(2\pi\sqrt{-1} \operatorname{Re} \int_{\gamma} \omega_{\alpha}\right),$$

independent of the choice of γ . Thus D determines a \dagger character on the 1-dimensional homology group, called the **integral character**. Conversely, every character on the homology group is shown to be the integral character of some D . In terms of this notion, Abel's theorem can be stated as follows: D is the principal divisor class if and only if $\chi_{\alpha}(D) = 1$ for every α . This result shows that the 1-dimensional homology group with integral coefficients and the group of the divisor classes of degree zero (with compact topology) are, with respect to integral characters, mutually dual (in the sense of Pontryagin) topological Abelian groups (\rightarrow 422 Topological Abelian Groups). For the relationship between Abelian integrals and Jacobian varieties, in particular the \dagger Jacobi inverse problem, \dagger Abelian functions, and \dagger Riemann theta functions, \rightarrow 3 Abelian Varieties L. Also \rightarrow references to 234 Kleinian Groups, 352 Quasiconformal Mappings, 367 Riemann Surfaces, 416 Teichmüller Spaces.

References

- [1] B. Riemann, *Gesammelte mathematische Werke und wissenschaftlicher Nachlass*, Teubner, 1876 (Collected works of B. Riemann, Dover, 1953).
 - [2] P. E. Appell and E. Goursat, *Théorie des fonctions algébriques et de leurs intégrales*, Gauthier-Villars, second edition, 1929.
 - [3] R. Dedekind and H. Weber, *Theorie der algebraischen Funktionen einer Veränderlichen*, J. Reine Angew. Math., 92 (1882), 181-290. (Gesammelte math. Werke, F. Vieweg, 1930, vol. 1, p. 2388350.)
 - [4] K. Hensel and G. Landsberg, *Theorie der algebraischen Funktionen einer Variablen und ihre Anwendung auf algebraische Kurven und Abelsche Integrale*, Teubner, 1902 (Chelsea, 1965).
 - [5] W. F. Osgood, *Lehrbuch der Funktionentheorie II*, Teubner, 1932.
 - [6] H. Weyl, *Die Idee der Riemannschen Fläche*, Teubner, 1913, third edition, 1955; English translation, *The concept of a Riemann surface*, Addison-Wesley, 1964.
 - [7] M. Eichler, *Einführung in die Theorie der algebraischen Zahlen und Funktionen*, Birkhäuser, 1963.
 - [8] C. Chevalley, *Introduction to the theory of algebraic functions of one variable*, Amer. Math. Soc. Math. Surveys, 1951.
 - [9] C. Siegel, *Topics in complex function theory II*, Wiley-Interscience, 1971.
- For history and further references,
- [10] A. Brill and M. Noether, *Die Entwicklung der Theorie der algebraischen Funktionen*, Jber. Deutsch. Math. Verein., 3 (1894), 107-566.
- For the classical theory of algebraic functions of two variables,
- [11] E. Picard and G. Simart, *Théorie des fonctions algébriques de deux variables indépendentes*, Gauthier-Villars, I, 1897; II, 1906.

12 (VIII.1) Algebraic Geometry

A. Introduction

Algebraic geometry is the branch of mathematics that deals with \dagger algebraic varieties, that is, point sets defined by several algebraic equations in a space of any dimension or those derived from these sets by means of certain constructions (\rightarrow 16 Algebraic Varieties). It may also be considered to be a theory of the \dagger field of algebraic functions in several variables in geometric language, and it is closely related to the theories of complex analytic manifolds, commutative algebra, and homological algebra. It also has an important connection with number theory through the theories of automorphic functions, Diophantine equations, and zeta functions.

To investigate local properties of algebraic varieties we consider varieties embedded in a \dagger affine space; to study global properties we usually consider varieties contained in \dagger projective spaces. A quantity (or property) that is invariant under \dagger projective transformations, isomorphisms, i.e., \dagger biregular and \dagger birational transformations, or birational transformations is called a \dagger projective invariant, a **relative invariant**, or an **absolute invariant** (birational invariant), respectively. The study of projective invariants is a part of projective geometry, whose methods are important in algebraic geometry. The notions of relative invariant and absolute invariant are used, for example, in the classification of algebraic varieties.

We usually assume that the coordinates of each point of the variety belong to a certain fixed field K . In the classical case, namely, when the field K is the field \mathbb{C} of complex numbers, the algebraic varieties are considered as complex spaces and are studied by applying the theories of partial differential equations, differential geometry, etc. Topological methods may also be applied. Algebraic geometry originated from such studies, but, for the study of properties such as rational mappings or algebraic systems, it became necessary to consider as well the case where the ground field K is not algebraically closed. Furthermore, to apply this to number theory, it is necessary to establish the theory over the field of any characteristic p . For this purpose it is necessary to establish a theory for varieties having ground domains as general as possible.

B. History

Analytic geometry began with the study of lines and quadratic curves (surfaces) and later came to include the study of cubic and quartic curves (surfaces), and so on. These subjects originally belonged exclusively to analytic (or projective) geometry. At that time, the study could not have been described by so specific a title as algebraic geometry.

The study of such theories as the construction of an algebraic plane curve by families of curves of lower degree or the algebraic m - n correspondence on a straight line probably began with research such as that by M. Chasles. The most outstanding event in the history of algebraic geometry was the introduction and development of the theory of algebraic functions (- 11 Algebraic Functions) by B. Riemann (1857). Before that time the degree of an algebraic curve (surface) was the only quantity known to be a projective invariant of the curve (surface).

With the theory of algebraic functions, Riemann gathered into one family all the curves that can be transformed onto each other by birational transformations. As the basis for his study, Riemann examined birational transformations in place of projective transformations. This idea led to the notion of the so-called Riemann surface. The genus of the surface was obtained as the characteristic number of the family of curves. The concept of genus was the first absolute invariant to appear in the history of algebraic geometry. Riemann based his theory on Abelian integrals using Dirichlet's principle, under the assumption that any algebraic curve reduces to one without singularities.

After Riemann many mathematicians tried

to reconstruct the theory more precisely without using transcendental methods. M. Noether attempted this reconstruction by using geometric methods. Using the Cremona transformation, he confirmed Riemann's assumption for curves: that any algebraic curve on a plane can be transformed by a birational transformation to a plane curve without singularities except for simple nodes. He also contributed in making more precise the basis conditions for the Riemann-Roch theorem, which is considered to be one of the most important theorems in the field. His results on space curves and surfaces are also noteworthy. J. Plücker defined the concept of genus in geometric terms and introduced the Plücker coordinates. A. Cayley and A. Brill worked along similar lines. Cayley's idea was developed later by B. L. van der Waerden and W. L. Chow, who introduced the associated form of an algebraic variety and its Chow coordinates.

Around 1890 the Italian school of algebraic geometry appeared. Following the tradition established by Noether, they employed algebraic-geometric methods and uncovered many new facts concerning algebraic surfaces. Among those who belonged to this school were G. Castelnuovo, F. Enriques, and F. Severi.

In France, H. Poincaré and E. Picard initiated their study of algebraic functions of two complex variables. After them S. Lefschetz investigated the theory of complex algebraic surfaces [11, 123]. The results attained by the Italian and French schools were very suggestive but lacked rigorous foundations.

On the other hand, rigorous number-theoretic theories of algebraic curves appeared in Germany. R. Dedekind and H. Weber developed the theory of algebraic function fields parallel to that of algebraic number fields. K. Hensel introduced the concept of p -adic numbers in analogy to power series expansions of analytic functions. E. Noether constructed an abstract theory of polynomial ideals from a formal theory by E. Lasker and F. S. Macaulay. Under her influence there appeared the arithmetic algebraic geometry (of curves) over an abstract field as developed by F. K. Schmidt and others.

In the higher-dimensional case, van der Waerden attempted to create a more rigorous foundation for algebraic geometry under the influence of Noether's abstract ideal theory (c. 1930) [14]. He introduced the concept of generic points and specialization, and specifically defined the multiplicity of intersections of two varieties in a projective space. He succeeded in getting a rigorous proof of Bezout's theorem: In n -dimensional projective

space, the number of intersections of an r -dimensional algebraic subvariety of degree l with an $(n - r)$ -dimensional subvariety of degree m is always lm if they intersect in only a finite number of points.

The problem of intersections was taken up by C. Chevalley and A. Weil in the 1940s. Chevalley developed the ideal theory of \dagger local rings (studied initially by W. Krull); he introduced topological concepts and applied them to the problem of intersections. The theory in this direction was later extended further by P. Samuel, M. Nagata, and J.-P. Serre.

Weil gave foundations of algebraic geometry over an abstract field and reconstructed the theory by introducing geometric language to designate objects of abstract algebra [15]. He thus gave quite a new aspect to the theory and extended H. Hasse's arithmetization of the theory of algebraic functions in one variable to the case of several variables. Reconstructing Severi's theory of algebraic correspondence over abstract fields, he succeeded in proving an analogy of the \dagger Riemann hypothesis on \dagger congruent zeta functions (- 450 Zeta Functions N). He also constructed, purely algebraically, the entire theory of \dagger Abelian varieties independent of characteristic.

Around 1930, O. Zariski gave another foundation to algebraic geometry by applying the generalized \dagger Valuation theory that had been introduced by Krull. Zariski clarified especially the properties of birational transformations by using valuation theory. Zariski's main theorem states that if a birational mapping is not \dagger regular at a \dagger normal point P (- 16 Algebraic Varieties 1), each component of the image of P by the mapping is of dimension ≥ 1 .

Zariski also solved the problem of \dagger resolution of singularities in the affirmative in the case of characteristic 0 for varieties of dimension ≤ 3 . The affirmative resolution of this problem (which Riemann assumed) says that any algebraic variety in a projective space can be transformed birationally to a projective algebraic variety without singularities. In 1964, H. Hironaka gave an affirmative answer for any dimension in the case of characteristic 0.

Along with the achievements in algebraic methods, great development took place in analytic methods. Unification of the concepts of Riemann surfaces and \dagger Riemannian manifolds led to the concept of \dagger complex analytic manifolds. Furthermore, G. de Rham's theorem on the duality of topologically defined homology and cohomology based on differential forms was proved; also, W. V. D. Hodge's theory of \dagger harmonic integrals was developed. In the case of the complex dimension 1, any

compact Riemann surface is derived from a certain projective algebraic curve. However, the situation is not so simple in the case of higher dimensions. Weil's concept of \dagger abstract complete algebraic varieties can be considered as an analog of compact complex manifolds. If a compact complex analytic variety is projective, then it must be an algebraic variety (\dagger Chow's theorem). K. Kodaira proved that a necessary and sufficient condition for a compact complex analytic manifold to be biholomorphically equivalent to a projective complex analytic manifold is that the manifold is a \dagger Hodge variety.

Using results on harmonic integrals, J. Igusa and Weil established the theory of \dagger Picard and \dagger Albanese varieties associated with algebraic manifolds of arbitrary dimensions as a generalization of the theory of \dagger Jacobian varieties associated with algebraic curves (- 9 Algebraic Curves E). Thus many ambiguities in the theory as developed by the Italian school were clarified. Later the theory was generalized to the case of characteristic p by T. Matsusaka, Chow, and S. Lang. The duality theorem in this case was later proved by M. Nishi and P. Cartier (- 3 Abelian Varieties D).

The concept of \dagger sheaves (- 383 Sheaves) had already been used in Kodaira's theory. Serre defined an abstract algebraic variety as a \dagger ringed space by using an analogy to the concept of complex analytic spaces; he considered it as a topological space with respect to the \dagger Zariski topology. By introducing \dagger coherent algebraic sheaves, Serre clarified the idea that classical invariants (such as \dagger arithmetic genus) may be considered cohomological quantities (\rightarrow 16 Algebraic Varieties E).

A. Grothendieck invented the concept of a \dagger scheme, which is far more general than that of an algebraic variety, by admitting the existence of \dagger nilpotent elements in structure sheaves and taking as a coordinate ring a general commutative ring with unity element. By the device of taking into account nilpotent elements, an analog of the method of successive approximation in analysis is now applicable. By masterful use of cohomological techniques, Grothendieck derived many results, including Zariski's important theorems.

In the classical case, F. Hirzebruch generalized the Riemann-Roch theorem to higher-dimensional manifolds. He made use of the language of sheaves and some topological results of A. Borel and R. Thom [S]. Later Grothendieck generalized the theorem for the abstract case as well. His idea in this work is recognized as the origin of \dagger K-theory.

Every nonsingular complete curve of genus 0 is isomorphic to the projective line, and any

nonsingular complete curve of genus 1 is isomorphic to a projective curve defined by the equation $X_2^2 X_0 = X_1(X_1 - X_0)(X_0 - \lambda X_1)$ for some $\lambda \neq 0, 1$. On the other hand, the set of all isomorphism classes of nonsingular complete curves of genus $g > 1$ is parametrized by a normal quasiprojective variety of dimension $3g - 3$. Such facts were first discussed by Riemann as the problem of moduli.

Concerning the moduli of manifolds of higher dimensions in the classical case, Kodaira and D. C. Spencer developed their theory of deformations of complex structures (\rightarrow 72 Complex Manifolds G). The meaning of number of moduli is clarified by deformation theory. Deformation theory has been extended in various ways, and deformation is considered as one of the fundamental concepts in algebraic geometry.

To investigate the global structure of the moduli varieties, D. Mumford introduced geometric invariant theory (- 16 Algebraic Varieties W).

Étale and crystalline cohomologies initiated by Grothendieck and others are useful for the study of algebraic varieties of positive characteristic. In particular, the conjecture made by Weil concerning congruent zeta functions has been solved affirmatively by P. Deligne with the help of étale cohomology [2].

Many important questions have been answered by means of the geometric theory. For example, (1) every vector bundle on \mathbf{A}_k^n is trivial (- 16 Algebraic Varieties Z); (2) there exist unirational but nonrational fields over \mathbf{C} (- 16 Algebraic Varieties J); (3) the fundamental groups of the complements of node-curves on $\mathbf{P}_\mathbf{C}^2$ are commutative (- 16 Algebraic Varieties I); (4) the cancellation theorem holds for \mathbf{A}_k^2 (\rightarrow 15 Algebraic Surfaces H); (5) whenever the tangent vector bundles are ample, the varieties are \mathbf{P}_k^n (\rightarrow 16 Algebraic Varieties R).

References

- [1] G. Castelnuovo, *Sulle funzioni Abelianne*, Memorie scelte, Zanichelli, 1937, 5299549.
- [2] P. Deligne, *La conjecture de Weil I*, Publ. Math. Inst. HES, 43 (1974), 2733307.
- [3] J. Dieudonné, *Cours de Géométrie Algébrique, I. Aperçu Historique sur le Développement de la Géométrie Algébrique*, Presses Univ. France, Collection Sup., 1974.
- [4] F. Enriques, *Memorie scelte di geometria*, Zanichelli, I, 1956; II, 1959; III, 1966.
- [5] A. Grothendieck, *The cohomology theory of abstract algebraic varieties*, Proc. Intern. Congr. Math., Edinburgh, 1958, Cambridge Univ. Press, 1958, 103-118.

- [6] A. Grothendieck, *Éléments de géométrie algébrique*, Publ. Math. Inst. HES, 4 (1960); 8 (1961); 11 (1961); 17 (1963); 20 (1964); 24 (1965); 28 (1966); 32 (1967).
- [7] H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero*, Ann. Math., (2) 79 (1964), 1099326.
- [8] F. Hirzebruch, *Topological methods in algebraic geometry*, Erg. Math., Springer, third edition, 1966.
- [9] K. Kodaira, *Collected Works I, II, III*, Iwanami, 1975.
- [10] S. Lefschetz, *L'analyse situs et la géométrie algébrique*, Gauthier-Villars, 1924.
- [11] E. Picard and G. Simart, *Théorie des fonctions algébriques de deux variables indépendantes*, Gauthier-Villars, I, 1897; II, 1906.
- [12] D. Quillen, *Projective modules over polynomial rings*, Inventiones Math., 36 (1976), 1677171.
- [13] J. G. Semple and L. Roth, *Introduction to algebraic geometry*, Clarendon Press, 1949.
- [14] B. L. van der Waerden, *Einführung in die algebraischen Geometrie*, Springer, 1939.
- [15] A. Weil, *Foundations of algebraic geometry*, Amer. Math. Soc. Colloq. Publ., 1946, 1962.
- [16] A. Weil, *Oeuvres Scientifiques, Collected Papers I, II, III*, Springer, 1980.
- [17] O. Zariski, *Collected papers*, MIT Press, I, 1972; II, 1973; III, 1978; IV, 1979.

13 (IV.1 1) Algebraic Groups

A. Definitions and General Remarks [1, 2, 8]

Let k be a field and Ω a universal domain containing it. An **affine algebraic group** G defined over k is, by definition, a group G which has the structure of a (not necessarily irreducible) algebraic variety defined over k in an affine space Ω^N such that the group operation $(x, y) \mapsto x^{-1}y$ on G is an everywhere regular mapping defined over k . For such a group G , the set G_k of all k -rational points on G is an abstract group. The irreducible component G_0 of G (viewed as an algebraic set) containing the identity element e is unique and is a normal subgroup defined over k with finite index in G ; the decomposition of G into (absolute) irreducible components coincides with that into the cosets of G by G_0 . When $G = G_0$, the group G is called **connected**. It should be noted that for an algebraic group G defined over k , a coset gG_0 is not necessarily defined over k ; and if it is, a representative g cannot necessarily be taken to be k -rational.

However, if k is an infinite field and G is connected, and if moreover k is t perfect or G is t reductive, then G_k is t Zariski dense in G ([9], A. Borel and T. A. Springer [30]); hence G is uniquely determined by G_k . (In this case, G_k is sometimes referred to as a **k-group**.) When k is a t topological field, G_k is a ${}^+$ topological group with respect to the natural topology defined by that of k , which is generally stronger than the Zariski topology on G_k . For instance, when $k = \mathbf{R}$, $(G_0)_{\mathbf{R}}$ is a t Lie group with finitely many connected components.

A subgroup H of an affine algebraic group G , which is Zariski closed, is an affine algebraic group with respect to its natural induced structure and is called an **algebraic subgroup** of G . If H is defined over k , then H is k -closed; the converse is also true when k is t perfect. (An affine algebraic set A is called **k-closed** if A is a common zero of a set of polynomial equations with coefficients in k . The set A is k -closed if and only if all irreducible components of A are defined over the algebraic closure \bar{k} of k , and for every t Galois automorphism σ of k/\bar{k} , $A^\sigma = A$.) The notions of homomorphism and isomorphism for algebraic groups can be defined in a natural manner. For instance, for affine algebraic groups G and G' defined over k , a **rational homomorphism** $\varphi: G \rightarrow G'$ defined over k , or a **k-morphism** for short, is a homomorphism of G into G' that is at the same time an (everywhere regular) rational mapping defined over k . For a k -morphism φ of G into G' , the (set-theoretic) image $\varphi(G)$ is a closed subgroup of G' defined over k , the kernel $\varphi^{-1}(e')$ is a k -closed subgroup of G , and $\dim \varphi(G) = \dim G - \dim \varphi^{-1}(e')$. In particular, when $\dim G = \dim \varphi(G) = \dim G'$ (or equivalently, when $\varphi(G_0) = G'_0$ and $\varphi^{-1}(e')$ is finite), φ is called an **isogeny**. (Two groups G and G' are called **isogenous** if there exist a third group G'' and isogenies $G'' \rightarrow G$, $G'' \rightarrow G'$.) When a k -morphism φ is bijective and φ^{-1} is also an (everywhere regular) rational mapping defined over k , φ is called a **birational isomorphism** defined over k , or a **k-isomorphism** for short. It should be noted that a rational homomorphism which is an isomorphism of abstract groups is not necessarily an isomorphism of algebraic groups (e.g., a Frobenius homomorphism); a similar statement holds for an injective continuous homomorphism of topological groups.

Given a connected affine algebraic group G and a closed subgroup H , both defined over k , the quotient space G/H has the uniquely determined structure of an algebraic variety defined over k such that the canonical mapping $G \rightarrow G/H$ is separable. The t function field of G/H is then identified with the subfield of the function field of G formed by all H -invariant ele-

ments. In particular, if H is a closed normal subgroup, then G/H has a natural structure of an affine algebraic group (defined over the same ground field k) [8, 10, 11].

As an example, we have the group $GL(n)$ of all $n \times n$ nonsingular matrices (x_{ij}) . $GL(n)$ may be viewed as an algebraic set in Ω^{n^2+1} , defined by a single equation $\det(x_{ij}) \neq 0$, and as such is a connected algebraic group defined over the t prime field. In general, an algebraic group reahzed as a closed subgroup of $GL(n)$ is called a **linear algebraic group**. Since an affine algebraic group is always isomorphic to a linear algebraic group, these two terminologies are essentially synonymous [1, 2].

B. Generalization of the Definition

Replacing the term *affine algebraic set* (or *affine variety*) in the definition of an affine algebraic group by a more general term t *algebraic variety*, we obtain the notion of an **algebraic group** (**algebraic group variety**, or simply **group variety**). On this subject, the following facts are fundamental. A t complete connected algebraic group is an Abelian variety (\rightarrow 3 Abelian Varieties). More generally, given a connected algebraic group G defined over k , there always exists a (k -closed) largest linear connected closed normal subgroup L , and the factor group G/L is an Abelian variety. Furthermore, for a closed normal subgroup H of G , the factor group G/H is complete if and only if $H \supset L$ (**Chevalley's theorem** [8]). In particular, if a connected algebraic group G is complete and linear at the same time, then G reduces to the identity group. In view of these theorems, the study of algebraic groups can be reduced, in a sense, to the study of Abelian varieties and linear algebraic groups. For this reason, we henceforth restrict ourselves to linear algebraic groups, which are simply called algebraic groups. (The notion of t generalized Jacobian variety, introduced by M. Rosenlicht [12], is an example of an algebraic group in a general sense; \rightarrow 9 Algebraic Curves.) The notion of algebraic groups has been generalized further to that of t group schemes by A. Grothendieck [3, 4].

C. Lie Algebras

Since an algebraic group G defined over k has no singularities, the t tangent space \mathfrak{g} to G at the identity element e is defined and has the same dimension as G : $\dim \mathfrak{g} = \dim G$. The space \mathfrak{g} can be identified in a natural manner with the space of all left-invariant t derivations of the function field of G_0 and thus has the struc-

ture of a Lie algebra defined over k (\rightarrow 248 Lie Algebras). We call \mathfrak{g} (the Lie algebra \mathfrak{g}_k over k of all k -rational points in \mathfrak{g}) the **Lie algebra** of G (of the k -group G_k). If G is a linear algebraic group contained in $GL(n)$, then \mathfrak{g}_k is a Lie subalgebra of $\mathfrak{gl}(n, k)$ with the Lie product defined by $[x, y] = xy - yx$; a linear Lie algebra corresponding to a linear algebraic group is called an **algebraic Lie algebra**. When the characteristic of k is zero, conditions for a linear Lie algebra to be algebraic can be given in terms of the **replica** [5]. Also, in the case of characteristic zero, for $x \in \mathfrak{gl}(n, k)$, $x \in \mathfrak{g}_k$ if and only if $\exp(tx) \in G$, where t is a variable over k and $\exp(tx)$ is understood as a \dagger formal power series in t (contained in Ω). From this, we can prove, exactly as in the theory of Lie groups (\rightarrow 249 Lie Groups), a one-to-one correspondence between k -closed subgroups H of G and algebraic Lie subalgebras \mathfrak{h}_k of \mathfrak{g}_k , establishing a complete parallelism of the theories of algebraic groups and Lie algebras [5]. This parallelism breaks down when k has positive characteristic [1, 2]. On the other hand, over a field of characteristic $p > 0$, we have **formal groups**, an analog of local Lie groups introduced by J. Dieudonné [13], and also \dagger hyperalgebras, which play the role of Lie algebras in characteristic 0.

D. Tori [1, 2]

The group $G_m = GL(1)$, the multiplicative group of nonzero elements in Ω , is a 1-dimensional connected algebraic group defined over the prime field. In general, an algebraic group G that is isomorphic to the direct product $(G_m)^n$ is called an **(algebraic) torus**. When a torus G defined over k is isomorphic to (G_m) over an extension K of k , G is called **K-trivial** (or **K-split**), and the field K is called a **splitting field** for G . A torus G defined over k always has a splitting field K which is a finite separable extension of k .

In general, a rational homomorphism χ of an algebraic group G into G_m is called a **character** of G . If we define the sum of two characters χ_1 and χ_2 of G by $(\chi_1 + \chi_2)(g) = \chi_1(g) \cdot \chi_2(g)$ ($g \in G$), the totality of characters of G is an additive group, called the **character module** of G and denoted by $X(G)$. Let G be a torus defined over k and $X = X(G)$ its character module, and let K be a splitting field for G that is a finite \dagger Galois extension of k . If a K -isomorphism $G \cong (G_m)^n$ is given by the correspondence $G \ni g \rightarrow (\chi_1(g), \dots, \chi_n(g))$, then the χ_i are characters of G , and X is a \dagger free module of rank n generated by χ_1, \dots, χ_n . Furthermore, if Γ denotes the Galois group of K/k , then for $\sigma \in \Gamma$ and $\chi \in X$, the conjugate χ^σ is also a char-

acter of G ; under this action of Γ , X becomes a right f -module. We have complete duality between a torus G and its character module X in the following sense. There exists a one-to-one correspondence between the closed subgroups G_1 (defined over k) of G and a $(\Gamma$ -invariant) submodule X_1 of X for which X/X_1 has no p - \dagger torsion (where p is the characteristic of k). This correspondence is determined by the relation of the annihilators $X_1 = G_1^\perp$, $G_1 = X_1^\perp$, and under the correspondence the character modules of G_1 and of G/G_1 are canonically identified with X/X_1 and X_1 , respectively. Furthermore, let G' be another torus (defined over k and split over K) with the character module X' , and suppose that we have a $(k$ -) homomorphism $\varphi: G \rightarrow G'$. Then we can define a $(\Gamma$ -) homomorphism ${}^t\varphi: X' \rightarrow X$, called the **dual homomorphism** of φ , by the relation ${}^t\varphi(\chi') = \chi' \circ \varphi$ for $\chi' \in X'$; conversely, any $(\Gamma$ -) homomorphism of X' into X is obtained uniquely in this manner. In particular, φ is a $(k$ -) isomorphism if and only if its dual ${}^t\varphi$ is a $(\Gamma$ -) isomorphism. Since for any free $(\Gamma$ -) module X of finite rank there always exists a torus G (defined over k and split over K) such that $X(G) \cong X$ (as a special case of the existence theorem of k -forms), the \dagger categories of all tori (defined over k and split over K) and that of all free $(\Gamma$ -) modules of finite rank are mutually dual.

E. Semisimple Elements and Unipotent Elements

A matrix a is called \dagger semisimple if it is diagonalizable, i.e., if the \dagger minimal polynomial of a has only simple roots. A matrix a is called \dagger unipotent if $a - 1$ is nilpotent, i.e., if all characteristic roots of a are equal to 1. (When the characteristic of the ground field is zero, the unipotent elements u in $GL(n, k)$ and the nilpotent elements x in $\mathfrak{gl}(n, k)$ are in one-to-one correspondence by the relation $u = \exp x$.) Any nonsingular matrix a can be written uniquely as a product of a nonsingular semisimple matrix a' and a unipotent matrix a'' of the same size which are mutually commutative: $a = a' a'' = a'' a'$ (\dagger multiplicative Jordan decomposition); a' (a'') is called the **semisimple (unipotent) part** of a and is denoted by a_s (a_u); a_s can be expressed as a polynomial of the matrix a with scalar coefficients. For an element a of a (linear) algebraic group G , the semisimplicity (unipotency) of a does not depend on the matrix representation of G . Moreover, these properties are preserved by homomorphisms of algebraic groups. Also, if $a \in G$, then $a, a_u \in G$.

For an algebraic group G , we denote the

totality of semisimple (unipotent) elements contained in G by $G_s (G_u)$ and call it the **semi-simple (unipotent) part** of G . (Note that G_s and G_u are not necessarily subgroups.) A torus G is then characterized by the property that $G = G_0 = G_s$. On the other hand, an algebraic group G such that $G = G_u$ is called **unipotent**. For instance, the additive group of the universal domain,

$$G_a \cong \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \middle| x \in \Omega \right\},$$

is a 1-dimensional connected unipotent algebraic group.

F. Solvable Groups and Nilpotent Groups [1, 2, 7, 9]

For two closed normal subgroups H_1, H_2 (defined over k) of an algebraic group G , the \dagger commutator group $[H_1, H_2]$ (in the sense of abstract group theory) is also a closed normal subgroup (defined over k) of G . In view of this fact, an algebraic group G is called **solvable (nilpotent)**, when it is \dagger solvable (\dagger nilpotent) as an abstract group. For example, the totality $T(n)$ of $n \times n$ nonsingular *upper unipotent* matrices, i.e., matrices of the form

$\begin{pmatrix} * & \dots & * \\ & \ddots & \\ 0 & & * \end{pmatrix}$, is a connected solvable algebraic group. A unipotent algebraic group is always nilpotent.

For any connected solvable algebraic group $G \subset GL(n)$, there always exists an element a in $GL(n)$ such that $a^{-1}Ga \subset T(n)$ (**Lie-Kolchin theorem** [1, 2]). A connected solvable algebraic group G has a \dagger composition series $G = G_0 \supset G_1 \supset \dots \supset G_r = \{e\}$ such that each G_i is a connected closed normal subgroup of G and G_{i-1}/G_i is isomorphic to either G_m or G_s . If G is defined over k , the subgroup G_u is a connected k -closed normal subgroup of G , and for any maximal torus T in G , we have a decomposition into a \dagger semidirect product $G = T \cdot G_u$ (in the sense of algebraic groups, i.e., the natural map $T \times G_u \rightarrow G$ is birational). It is known that for any algebraic group G defined over k there exists a maximal torus defined over k ([4], A. Borel and T. A. Springer [30]). G is nilpotent if and only if G has a unique maximal torus T ; when that is so, $T = G_s$ and T is contained in the \dagger center of G . For a connected solvable algebraic group G defined over k , we can take $a \in GL(n, k)$ such that $a^{-1}Ga \subset T(n)$ (see the Lie-Kolchin theorem) if and only if all characters $\chi \in X(G)$ are defined over k ; when this condition is satisfied, G is called **k -solvable**. G_u is then defined over k , and G/G_u is a k -trivial torus.

When the characteristic of k is zero, any commutative unipotent algebraic group (defined over k) is (k -) isomorphic to the direct product $(G_a)^n$. When k is an algebraically closed field of characteristic $p > 0$, any connected commutative unipotent algebraic group defined over k is k -isogenous to a direct product of a certain number of the groups W_m of \dagger Witt vectors (of length m) (Chevalley-Chow theorem [12]). A 1-dimensional connected unipotent algebraic group defined over a perfect field k is k -isomorphic to G_a [1, 2].

G. Borel's Theory

Let G be an algebraic group and V an algebraic variety (both defined over k). We say that V is a **transformation space** of G (defined over k), or simply G acts on V , if there is given an everywhere regular rational mapping $G \times V \ni (g, v) \rightarrow gv \in V$ (defined over k) such that $g_1(g_2v) = (g_1g_2)v$, $ev = v$ ($g_1, g_2 \in G, v \in V$). When the action of G on V is \dagger transitive, V is called a \dagger homogeneous space of G . For a closed subgroup H of a connected algebraic group G (both defined over k), the quotient space G/H has the natural structure of a homogeneous space of G (defined over k). A. Borel [1, 2] proved the following theorems:

(1) If G is a connected solvable algebraic group and V a complete transformation space of G , then G has at least one fixed point in V . More precisely, in order that a connected algebraic group G defined over k be k -solvable, it is necessary and sufficient that for any complete transformation space V of G defined over k for which $V_k \neq \emptyset$, G have at least one k -rational fixed point in V [9].

(2) Let G be a connected algebraic group. A maximal connected solvable closed subgroup of G is called a **Borel subgroup** of G . Then (i) all pairs (T, B) formed by a maximal torus T in G and a Borel subgroup B containing it are conjugate to each other with respect to inner automorphisms of G . (ii) For a closed subgroup H of G , the quotient space G/H is complete if and only if H contains a Borel subgroup of G ; and, when that is so, G/H is actually a \dagger projective algebraic variety. (For instance, if $G = GL(n)$, $B = T(n)$, then G/B is a so-called \dagger flag manifold.) (iii) The conjugates of $B(T)$ cover the whole group G (G_s). A closed subgroup of G is called **parabolic** if it contains a Borel subgroup of G . A parabolic subgroup H coincides with its own \dagger normalizer $N(H)$; in particular, H is always connected. Parabolic subgroups are significant in the theory of automorphic functions. (For the parabolic subgroups associated with BN-pairs \rightarrow Sections Q, R.)

When G is a connected algebraic group defined over a perfect field k , proposition (i) can be sharpened: The pairs (A, H) formed by a maximal k -trivial torus A in G and a maximal connected k -solvable subgroup H containing it are conjugate to each other with respect to the inner automorphisms defined by elements in G_k . The normalizer $N(H)$ of such a k -solvable subgroup H is a minimal k -closed parabolic subgroup of G . When the maximal connected k -solvable subgroups of G are reduced to the identity group, G is called **k -compact** or **k -anisotropic**. (Otherwise, G is called **k -isotropic**.) For instance, the † orthogonal group $G = SO(n, f)$ of a † quadratic form f of n variables is k -compact if and only if the form f is **anisotropic**, i.e., the homogeneous equation $f = 0$ has no solution other than zero in k . Similar facts hold for other classical groups. When k is a † local field, G is k -compact if and only if G_k is compact as a topological group. In general, a k -compact group is † reductive.

H. The Weyl Group

Let G be a connected algebraic group and Q an arbitrary torus in G . The † centralizer $Z(Q)$ of Q is then connected and coincides with the connected component of the normalizer $N(Q)$. Hence the factor group $W = N(Q)/Z(Q)$ is finite and can be identified with a subgroup of the automorphism group of Q (or of its character module $X(Q)$) in a natural manner. The group W is called the **Weyl group** of G relative to Q . In particular, when $Q = T$ (a maximal torus), the order of W is equal to the number of Borel subgroups containing T . In this case, the centralizer $C = Z(T)$ is called a **Cartan subgroup** of G ; it is characterized by the property that C is a (maximal) connected nilpotent closed subgroup of G which coincides with the connected component of its own normalizer $N(C)$. The notions of Borel subgroups, Cartan subgroups, and maximal tori are preserved under rational homomorphisms of algebraic groups.

1. Semisimple Groups and Reductive Groups

In an algebraic group G defined over k , there exists a largest connected solvable closed normal subgroup R , called the **radical** of G . The unipotent part R_u of R is called the **unipotent radical** of G . When $R = \{e\}$, G is called **semisimple**. When R is a torus, namely, $R_u = \{e\}$, G is called **reductive**. Semisimplicity and reductiveness are preserved under forming a direct product and taking the image (or inverse image) of an isogeny. For a reductive group G , the $+$ commutator subgroup $D(G)$ is

semisimple, and $G = D(G) \cdot R$, $D(G) \cap R = \text{finite}$; in other words, G is isogenous to the direct product of a connected semisimple algebraic group and a torus. In general, if R is the radical of a connected algebraic group G and R_u is the unipotent radical of R , then the factor groups G/R , G/R_u are semisimple and reductive, respectively. Furthermore, if the characteristic of the field k is zero, there exists a reductive closed subgroup H of G such that G decomposes into a semidirect product $G = H \cdot R_u$ (**Chevalley decomposition** [5]). (In this case, R and R_u are k -closed, and H can be taken to be k -rational; such an H is unique up to inner automorphisms defined by elements in G_k .) Also in the case of characteristic zero, reductive algebraic groups are characterized by the property that all rational representations are completely reducible. But when k has the characteristic $p > 0$, this property characterizes tori (M. Nagata).

J. Root Systems [1, 2, 14]

Let G be a connected semisimple algebraic group, T a maximal torus, and $X = X(T)$ its character module. A character $\alpha \in X$ is called a **root** of G relative to T if there exists an isomorphism x_α of G_a onto its image in G such that

$$t^{-1}x_\alpha(\xi)t = x_\alpha(\alpha(t)\xi) \quad \text{for all } \xi \in G_a, t \in T.$$

For a root α , such an isomorphism x_α is uniquely determined up to a scalar multiplication in G_a ; hence we put $P_\alpha = x_\alpha(G_a)$.

If we denote by r the totality of roots (relative to T), r satisfies the following axioms, where $E = X \otimes Q$ and E^* is the † dual space of E with respect to the inner product (\cdot, \cdot) : (i) For each $\alpha \in r$, there corresponds $\alpha^* \in E^*$ such that $\langle \alpha^*, \alpha \rangle = 2$ and $\langle \alpha^*, \beta \rangle \in \mathbb{Z}$ for all $\beta \in r$. (ii) If we define a reflection w_α of E by

$$w_\alpha x = x - \langle \alpha^*, x \rangle \alpha \quad \text{for } x \in E,$$

then $w_\alpha \beta \in r$ for all $\beta \in r$. (In particular, $w_\alpha \alpha = -\alpha \in r$.) (iii) If $\alpha, \beta \in r$ are linearly dependent, then $\beta = \pm \alpha$. (iv) If $\dim E = r$, r contains r linearly independent elements.

In general, a finite subset r in a finite-dimensional vector space E over Q satisfying the axioms (i)–(iv) is called a **root system** in E . (This root system is sometimes said to be reduced, to distinguish it from the root system defined in Section Q, which does not satisfy axiom (iii).) For a root system r , the elements α^* of E^* corresponding to $\alpha \in r$ are uniquely determined by these conditions, and the set $r^* = \{\alpha^*\}$ is a root system in E^* ($\alpha^* \in r^*$ is called a **coroot**). Also, the group W of linear transformations of E (E^*) generated by w_α

(w_{α}) with $\alpha \in \mathfrak{r}$ is finite and is called the **Weyl group** of the root system \mathfrak{r} . If we identify E^* with E by means of any W -invariant (positive definite) metric on E , then $a^* = (2/\langle \alpha, \alpha \rangle)\alpha$. When \mathfrak{r} is a root system of a semisimple algebraic group,

$$\langle \alpha^*, \chi \rangle \in \mathbb{Z} \quad \text{for all } \alpha \in \mathfrak{r}, \chi \in X, \quad (1)$$

so that X is W -invariant, and the Weyl group of the root system \mathfrak{r} can be identified with the group $N(T)/Z(T)$ of Section H. (In general, a maximal torus in a connected reductive algebraic group coincides with its own centralizer, so that the Weyl group W can be identified with $N(T)/Z(T)$.)

When a "linear ordering (compatible with the addition) is given in E , we denote by \mathfrak{r}_+ the set of all positive roots in \mathfrak{r} . An element $\alpha \in \mathfrak{r}_+$ is called a **simple root** if it cannot be written as $\alpha = \alpha' + \alpha''$ with $\alpha', \alpha'' \in \mathfrak{r}_+$. If $\Delta = \{\alpha_1, \dots, \alpha_r\}$ is the totality of (distinct) simple roots in \mathfrak{r}_+ , the elements $\alpha_1, \dots, \alpha_r$ are linearly independent, and any root $\alpha \in \mathfrak{r}$ can be written uniquely in the form $\alpha = \pm \sum_{i=1}^r m_i \alpha_i$, with $m_i \in \mathbb{Z}$, $m_i \geq 0$. In general, a subset A of \mathfrak{r} having this property is called a **fundamental system**; a fundamental system is always obtained in the manner explained from a linear ordering on E . For a fundamental system A , the cone A , in E^* , defined as the set of x in E^* satisfying the inequalities $\langle \alpha_i, x \rangle > 0$ ($1 \leq i \leq r$), is called a **Weyl chamber**. If we denote by L_α the hyperplane defined by the linear equation $\langle \alpha, x \rangle = 0$ for a root α , then $E^* = \bigcup_{\alpha \in \mathfrak{r}} L_\alpha = \bigcup_{\Delta} \Lambda_\Delta$, and W acts simply transitively on the set of all Weyl chambers $\{\Lambda_\Delta\}$. The Weyl group W is generated by \mathfrak{r} reflections w_{α_i} ($1 \leq i \leq r$).

In a semisimple algebraic group G , Borel subgroups B containing a (fixed) maximal torus T are in one-to-one correspondence with the fundamental systems A (or \mathfrak{r}_+) relative to T by the relation $B_u = \prod_{\alpha \in \mathfrak{r}_+} P_\alpha$, where $P_\alpha = x_\alpha(G_u)$. (More precisely, every element in B_u can be written uniquely as a product of the elements in P_α , where the ordering of the P_α is taken arbitrarily.)

K. Bruhat Decomposition

If we take a representative s_w of $w \in W$ in $N(T)$, there is a decomposition $G = \bigcup_{w \in W} B s_w B$ (disjoint union). Furthermore, if for $w \in W$ we put $N_w = \prod_{\alpha \in \mathfrak{r}_+ \cap w\mathfrak{r}_+} P_\alpha$, and in particular $N = N_e = B_u$, and denote by w_0 the unique element in W such that $w_0 \Delta = -\Delta$, then the element in $B s_w B$ can be written uniquely as a product of elements in $N_{w w_0}, s_w T, N$. Hence we have

$$G = \bigcup_{w \in W} N_{w w_0} \cdot s_w T \cdot N,$$

which is called a **Bruhat decomposition** of G . In

particular, if we put $N' = s_{w_0}^{-1} N s_{w_0}$ (which is the unipotent part of the Borel subgroup corresponding to $-\Delta$), then $N' T N$ is a Zariski open set in G , and the natural map $N' \times T \times N \rightarrow G$ is birational. This implies that the function field of G is rational (i.e., purely transcendental over Ω).

L. Structure of Semisimple Groups

A subset \mathfrak{r}_1 of a root system \mathfrak{r} is called a **closed subsystem** if $\mathfrak{r}_1 \cap \mathfrak{r} = \mathfrak{r}_1$, where \mathfrak{r}_1 denotes the submodule of X generated by \mathfrak{r}_1 . A closed subsystem satisfies conditions (i), (ii), and (iii) of a root system. For a closed subsystem \mathfrak{r}_1 of a root system \mathfrak{r} of a semisimple algebraic group G , the subgroup $G(\mathfrak{r}_1)$ of G generated by the P_α ($\alpha \in \mathfrak{r}_1$) is a semisimple closed subgroup with a maximal torus $T_1 = (G(\mathfrak{r}_1) \cap T)_0$, of which the root system relative to T_1 coincides with the restriction of \mathfrak{r}_1 on T_1 and the coroot system can be identified with $\mathfrak{r}_1^* = \{\alpha^* | \alpha \in \mathfrak{r}_1\}$. The subgroup $G(\mathfrak{r}_1)$ is normal if and only if $\mathfrak{r} - \mathfrak{r}_1$ is also a closed subsystem; when this is so, $G = G(\mathfrak{r}_1) \cdot G(\mathfrak{r} - \mathfrak{r}_1)$, $G(\mathfrak{r}_1) \cap G(\mathfrak{r} - \mathfrak{r}_1) = \text{finite}$. All connected closed normal subgroups of G are obtained in this manner. In order that G be **simple** (sometimes called **absolutely simple** or **almost simple**) as an algebraic group (i.e., without proper connected normal subgroups), it is necessary and sufficient that \mathfrak{r} be **irreducible** (i.e., \mathfrak{r} cannot be decomposed into a disjoint union of two proper closed subsystems). In general, a root system \mathfrak{r} can be decomposed uniquely into the disjoint union $\mathfrak{r} = \mathfrak{r}_1 \cup \dots \cup \mathfrak{r}_s$ of irreducible closed subsystems \mathfrak{r}_i such that $\mathfrak{r}_1 \cup \dots \cup \mathfrak{r}_i$ ($1 \leq i \leq s$) are also closed subsystems; correspondingly, G is isogenous to the direct product $G_1 \times \dots \times G_s$ of (absolutely) simple algebraic groups $G_i = G(\mathfrak{r}_i)$. (G is actually a direct product if it is simply connected or an adjoint group.) The subgroups G_i are determined uniquely and only by G .

M. k-Forms [15]

Let K be an extension of k and G_1 an algebraic group defined over K . An algebraic group G defined over k is called a **k-form** of G_1 if there is a K -isomorphism f of G onto G_1 . Suppose further that K/k is finite separable, and for every Galois automorphism σ of k/k , put $\varphi_\sigma = f^\sigma \circ f^{-1}$. Then φ_σ is an isomorphism of G_1 onto G_1^σ , and the φ_σ satisfy the relation $\varphi_\sigma^\tau \circ \varphi_\tau = \varphi_{\sigma\tau}$. Conversely, given a collection of isomorphisms $\{\varphi_\sigma\}$ satisfying these conditions, there always exists a k -form G (with a K -isomorphism f onto G_1 such that $\varphi_\sigma = f^\sigma \circ f^{-1}$), which is unique up to k -isomorphism (Weil). In particular, if K/k is a

finite Galois extension with Galois group Γ , then $\{\varphi_\sigma\}$ is a (continuous) $^+1$ -cocycle of Γ in $\text{Aut}_K(G_1)$ (the group of all K -automorphisms of G_1), and by the above correspondence the k -isomorphism classes of k -forms G are in one-to-one correspondence with the (continuous) 1 -cohomology classes of the cocycle $\{\varphi_\sigma\}$ (in the cohomology set $H^1(\Gamma, \text{Aut}_K(G_1))$) (- 172 Galois Theory J)

To a given finite separable extension K/k of degree d and an algebraic group G_1 defined over K of dimension n , we can associate a certain algebraic group $\mathfrak{R}_{K/k}(G_1)$ defined over k of dimension dn , which is obtained from G_1 by restricting the ground field [19]. A more precise definition is as follows. Let $\{\sigma_1, \sigma_2, \dots, \sigma_d\}$ ($\sigma_1 = 1$) be a set of automorphisms of \bar{k}/k such that $\sigma_i K$ ($1 \leq i \leq d$) are all distinct. Then one can find a k -form \tilde{G} of $\tilde{G}_1 = \prod_{i=1}^d G_1^{\sigma_i}$ with an isomorphism $\tilde{f}: \tilde{G} \rightarrow \tilde{G}_1$ such that $\tilde{\varphi}_\sigma = \tilde{f}^\sigma \circ \tilde{f}^{-1}$ is given by $\tilde{\varphi}_\sigma((x_i)) = (x_i \sigma)$, where i^σ is defined by the relation $(\sigma_i \sigma) K = \sigma_{i^\sigma} K$. If we denote by p_1 the canonical projection of \tilde{G}_1 onto its first component G_1 and put $p = p_1 \circ \tilde{f}$, then the pair (\tilde{G}, p) is uniquely characterized (up to k -isomorphism) by the following universality property: If \tilde{G}' is any algebraic group defined over k and φ is a K -morphism of \tilde{G}' into G_1 , then there exists a (uniquely determined) k -morphism $\tilde{\varphi}$ of \tilde{G}' into \tilde{G} such that $\varphi = p \circ \tilde{\varphi}$. The group \tilde{G} (together with p) is denoted by $\mathfrak{R}_{K/k}(G_1)$. For the group of rational points, $\tilde{G}_k = G_k$. When the algebraic group G_1 has some additional structure (such as that of $^+ \text{vector space}$, $^+ \text{algebra}$, etc.), then $\mathfrak{R}_{K/k}(G_1)$ automatically has the same kind of additional structure.

N. Chevalley's Fundamental Theorems

Let G, G' be connected semisimple algebraic groups, and let $T(T')$ be a maximal torus in $G(G')$, $X(X')$ its character module, $r(r')$ a root system of $G(G')$ relative to $T(T')$, etc. If we have an $^+ \text{isogeny}$ φ of G onto G' such that $\langle p(T) = T' \rangle$, then there is a bijection $\alpha \rightarrow \alpha'$ of r onto r' such that $\psi(\alpha') = q_\alpha \alpha$, where ψ is the dual homomorphism of φ , T and q_α is a positive integer, which equals 1 if the characteristic is zero and is a power of p if the characteristic is $p > 0$. Conversely, any injective homomorphism $\psi: X' \rightarrow X$ satisfying this condition (with respect to a certain bijection $r \rightarrow r'$ and q_α) comes from an isogeny $\varphi: G \rightarrow G'$ in the manner already stated. In particular, φ is an isomorphism if and only if ψ is an isomorphism such that $\psi(r') = r$ (i.e., $q_\alpha = 1$ for all $\alpha \in r$) [7]. The isomorphism class of G is thus completely determined by the pair (X, r) , so that we sometimes write $G = G(X, r)$. A connected semi-

simple algebraic group G defined over k is called of **Chevalley type** over k (or **k -split**) if there exists a k -trivial maximal torus T in G . If, in the above theorem, G and G' are of Chevalley type over k and T and T' are k -trivial, then the theorem remains true if we replace isogeny by k -isogeny. In particular, the k -isomorphism class of a connected semisimple algebraic group of Chevalley type over k is completely determined by (X, r) . Chevalley also showed that, for any pair (X, r) satisfying condition (1) above, there exists a connected semisimple algebraic group $G(X, r)$ of Chevalley type defined over the prime field. Therefore, since the classification of semisimple algebraic groups of Chevalley type is reduced essentially to that of root systems (X, r) , it turns out that, over any ground field k , there exist as many connected simple algebraic groups of Chevalley type as connected simple complex Lie groups (\rightarrow 249 Lie Groups; Appendix A, Table 5.1).

For a given semisimple algebraic group $G = G(X, r)$ defined over k , put $X_0 = r_z (=$ the submodule of X generated by r), $X^0 = \{x \in F | \langle \alpha^*, x \rangle \in \mathbb{Z} \text{ for all } \alpha \in r\}$. Then we have natural isogenies $G(X^0, r) \rightarrow G(X, r) \rightarrow G(X_0, r)$ (with $q_\alpha = 1$), all of which can be taken to be defined over k . The group $G(X^0, r)(G(X_0, r))$ is called the **simply connected group** (the **adjoint group**) isogenous to G . When the characteristic of k is zero, these isogenies (which are already known in the classical theory of complex Lie groups) are essentially the only possible isogenies among the semisimple algebraic groups. But when the characteristic is $p > 0$, there are, in addition to these, the Frobenius homomorphism (with $q_\alpha = p$) and the following "singular" isogenies (for which $q_\alpha = 1$ or p depending on α): $B_n \rightarrow C_n, F_4 \rightarrow F_4$ ($p = 2$), $G_2 \rightarrow G_2$ ($p = 3$). In particular, when k is a finite field, taking the set of fixed points of the singular k -isogenies, we obtain the simple finite groups of M. Suzuki and R. Ree (- 151 Finite Groups).

0. Classification Theory

A connected semisimple algebraic group G defined over k is called **k -(almost) simple** if there is no proper connected closed normal subgroup of G defined over k . (When G is k -simple and k -split, the factor group $D(G_k)/\text{center}$ is an abstract simple group except for a few special cases [17]. For more general results \rightarrow Section Q.) For a k -simple algebraic group G , let G_k be any one of its absolutely simple components, and let k be the smallest field of definition for G_k containing k . Then k_1/k is a finite separable extension, and G is k -isogenous to $\mathfrak{R}_{k_1/k}(G_k)$. Hence the problem of

classifying all k -simple groups (up to isogeny) is equivalent to that of finding all k -forms of simple groups of Chevalley type. This latter problem can be reduced, in principle, to the classification of compact k -forms and that of certain diagrams (i.e., † Dynkin diagrams along with an action of the Galois group) [16], J. Tits [30] (- Appendix A, Table 5.1). For instance, when k is a finite field (or, more generally, a field of dimension ≤ 1 [15]), there is no compact k -simple group; hence, using a simple classification theory of the diagrams, we can show that the only absolutely simple algebraic groups G defined over k are either of Chevalley type or of the types introduced by R. Steinberg (denoted by 2A_n , 2D_n , 3D_4 , 6D_4 , 2E_6). Connected semisimple algebraic groups composed of the groups of these types are characterized by the property that they have a Borel subgroup defined over k . Such groups are said to be of **Steinberg type over k** (or **k -quasi-split**). Absolutely simple algebraic groups over a p -adic field have been classified by M. Kneser and J. Tits [30, 31]. When the characteristic of k is not equal to 2, the classification of simple groups of classical type (except for the type D_4) is known to be equivalent to that of semisimple † associative algebras with † involution [18]. A similar relation also holds between some of the exceptional simple groups and † Cayley algebras or † Jordan algebras (H. Hijikata, T. A. Springer, J. Tits).

The following is a list of absolutely simple algebraic groups of classical type.

- I. k -forms of $SL(n)$ ($n \geq 2$).
 - 1.1. $G_{1k} = SL(m, \mathfrak{R}) = \{g \in M_m(\mathfrak{R}) \mid N(g) = 1\}$, where \mathfrak{R} is a † central division algebra over k with $(\mathfrak{R}:k) = r^2$, $n = mr$, and N denotes the † reduced norm in $M_m(\mathfrak{R})$.
 - 1.2. $G_{1k} = SU(m, \mathfrak{R}, f) = \{g \in SL(m, \mathfrak{R}) \mid f(gx, gy) = f(x, y) \text{ for } x, y \in \mathfrak{R}^m\}$, where \mathfrak{R} is a central division algebra over a quadratic extension k' of k with an involution ι of the second kind (which means that $\{\xi \in k' \mid \xi' = \xi\} = k$), $(K:k') = r^2$, $n = mr$, and f is a (nondegenerate) † Hermitian form of m variables over \mathfrak{R} with respect to the involution ι .
- II. k -forms of $SO(n)$ ($n \geq 3$, $n \neq 4$), $Sp(n)$ (n even, $n \geq 2$).

$G_{1k} = SU(n, \mathfrak{R}, f)$, where \mathfrak{R} is a central division algebra over k with an involution ι of the first kind (i.e., such that $\{\xi \in k \mid \xi' = \xi\} = k$), $(\mathfrak{R}:k) = r^2$, $n = mr$, and f is a nondegenerate c -Hermitian form of m variables over \mathfrak{R} with respect to the involution ι . In this case, $\dim\{\xi \in K \mid \xi' = \xi\} = r(r + \varepsilon_0)/2$ with $\varepsilon_0 = \pm 1$, and G_1 is a k -form of SO or Sp according as $\varepsilon\varepsilon_0 = 1$ or -1 . ($SO(8)$ may have other k -forms coming from the so-

called triality.) The case where the characteristic of k is 2 can also be discussed by a method given by J. Tits (*Inventiones Math.* 5, 1968).

When k is a local field or an algebraic number field, the only central division algebra with an involution of the first kind is a † quaternion algebra (and, if ι is the " † canonical involution," then $\varepsilon_0 = -1$).

P. Algebraic Groups over an Algebraic Number Field

Let G be a connected algebraic group defined over an algebraic number field k of finite degree. Let $\{v\}$ be the totality of † prime divisors (i.e., equivalence classes of valuations) of k . Taking the † restricted direct product of a family of locally compact topological groups $\{G_{k_v}\}$, we obtain a locally compact topological group G_A , called the **adele group** of G [19] (- 6 Adeles and Ideles). In particular, when $G = G_m$, the adele group $I = (G_m)_A$ is exactly the † idele group introduced by Chevalley in class field theory. If we identify $x \in G_k$ with an adele whose components are all equal to x , G_k becomes a discrete subgroup of G_A .

Concerning the finiteness property of G_A/G_k , the following results have been obtained [21, 22]: A character $\chi \in X_k(G)$ (= the module of all k -rational characters of G) gives rise to a (continuous) homomorphism $\chi_A: G_A \rightarrow I = (G_m)_A$. Put $G_A^0 = \{g \in G_A \mid |\chi_A(g)| = 1 \text{ for all } \chi \in X_k(G)\}$, where $||$ is the standard norm in I . Then G_A^0 is † unimodular, and the quotient space G_A^0/G_k is of finite volume with respect to the (unique) invariant measure on it. G_A^0/G_k (G_A/G_k) is compact if and only if the semisimple part G/R (the reductive part G/R_u) of G is k -compact. From the arithmetic point of view, it is important to determine explicitly the volume of G_A^0/G_k with respect to the invariant measure normalized in a certain manner; such a volume is called the **Tamagawa number** of G and is usually denoted by $\tau(G)$ [19, 23]. For instance, Siegel's formulas on the volume of the fundamental domain of the unit group of a quadratic form f over k are essentially equivalent to a theorem on the Tamagawa number stating that $\tau(SO(f)) = 2$.

Let \mathfrak{o} be the † ring of integers in k and L an \mathfrak{o} -lattice in the vector space on which G is acting. We can define in a natural manner an action of G_A on the set of all \mathfrak{o} -lattices; then the orbit $G_A L(G_k L)$ of L with respect to $G_A(G_k)$ is called the **genus (class)** of L . The † stability subgroup $G_{A,L}$ of L in G_A is open, and the double coset space $G_{A,L} \backslash G_A/G_k$ is finite (finiteness of the class number). Moreover, let $\{v_1, \dots, v_r\}$ be the totality of † infinite prime

divisors of k , and put $G_\infty = \prod_{i=1}^r G_{k_i}$. Then G_∞ is a Lie group, and the canonical projection G_L on G_∞ of $G_{A,L} \subset G_A$ is a discrete subgroup of finite type. (In general, (discrete) subgroups of G_∞ which are t commensurable with G_L are called **arithmetic subgroups**.) As in the adèle case, $\chi \in X_k(G)$ gives rise to a (continuous) homomorphism $\chi_\infty: G_\infty \rightarrow (\mathbf{R}^\times)'$, and if $G_\infty^0 = \{g \in G_\infty \mid \chi_\infty(g) = 1\}$, then the quotient space G_∞^0/G_L is of finite volume. Moreover, G_∞^0/G_L (G_∞/G_L) is compact if and only if G_A^0/G_k (G_A/G_k) is compact.

In addition to these, the t approximation theorem and the t Hasse principle are also extended to (classical, or general) algebraic groups (M. Eichler, M. Kneser, G. Shimura, Hijikata, Springer; \rightarrow [30]).

Q. Structure of Reductive Groups [24]

Let G be a connected t reductive group defined over a field k . Then G has a t maximal torus defined over k and G is t split over a finite t separable extension K of k . The structure of the group G_K can be discussed as in Sections I, J, K, and L. Here, we discuss the structure of G_k , i.e., a t k -form of G .

Let S be a **maximal k -split torus** of G , i.e., a k -subtorus of G that is k -split and maximal with respect to these properties. Any two such tori are conjugate over k , i.e., by an element of G_k . Their dimension is called the k -rank of G . To say that G is of k -rank zero is equivalent to saying that G is t k -anisotropic. The centralizer $Z = Z_\infty(S)$ of S in G is a reductive k -group, and its derived group is k -anisotropic. Let $N = N_\infty(S)$ be the normalizer of S in G . The t Weyl group ${}_k W = N/Z$ relative to S is called the **k -Weyl group** of G . A **k -root** of G with respect to S is a nontrivial character of S that appears when one diagonalizes the representation of S in the t Lie algebra \mathfrak{g} of G , S operating via adjoint representation. Denote by ${}_k r$ the set of all k -roots of G with respect to S .

There is a decomposition of the Lie algebra \mathfrak{g} of G : $\mathfrak{g} = \mathfrak{g}_0 + \sum_{\alpha \in {}_k r} \mathfrak{g}_\alpha$ where $\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \mid \text{Ad}(s)X = \alpha(s)X \text{ for all } s \in S\}$. Then \mathfrak{g}_0 is the Lie algebra of Z , and there is a unique unipotent k -subgroup P_α of G normalized by S such that its Lie algebra is \mathfrak{g}_α . The set ${}_k r$ is a t root system in a suitable Euclidean space whose Weyl group is isomorphic to ${}_k W$; if G is k -split, ${}_k r$ is the ordinary root system, and the P_α are as in Section J. In general, ${}_k r$ need not be reduced (i.e., axiom (iii) in Section J need not be satisfied), nor should $\dim \mathfrak{g}_\alpha = \dim P_\alpha$ always be 1.

A closed subgroup of G defined over k which is minimal among the parabolic subgroups of G is called a **minimal paraholic k -subgroup** of G . (If G is k -split, a minimal para-

bolic k -subgroup is a t Borel subgroup of G .) Any two minimal paraholic k -subgroups are conjugate to each other over k . If P is one such subgroup, then there exists a maximal k -split torus S such that P is the semidirect product of the reductive k -subgroup $Z = Z_\infty(S)$ and the t unipotent radical $U = R_\infty(P)$ of P . The expression $P = ZU$ is called a **Levi-decomposition** of P , and Z is called a **Levi-subgroup** of P . Any two Levi-subgroups of P are conjugate by an element in U_k . There is an ordering of ${}_k r$ such that P is generated by Z and P_α with $\alpha > 0$. The minimal parabolic k -subgroups containing a given maximal k -split torus S correspond to the t Weyl chambers of ${}_k r$. They are permuted simply transitively by the Weyl group ${}_k W$.

Fix an ordering of ${}_k r$, and let ${}_k \Delta$ be the t fundamental system of ${}_k r$ with respect to the given order. For any subset θ of ${}_k \Delta$, denote by P_θ the subgroup generated by $Z = Z_\infty(S)$ and P_α , where α is a linear combination of the roots of ${}_k r$ in which all roots not in θ occur with a coefficient ≥ 0 . Then P_θ contains P and, in particular, $P_\emptyset = P$. P_θ is called a **standard paraholic k -subgroup** of G containing P . Any paraholic k -subgroup of G is conjugate over k to a unique P_θ . If S_θ is the identity component of $\bigcap_{\alpha \in \theta} (\ker \alpha)$, then S_θ is a k -split torus of G and $P_\theta = Z(S_\theta)R_u(P_\theta)$. This shows that any paraholic k -subgroup of G has a Levi-decomposition and its Levi-subgroups are conjugate to each other over k . Let P be a minimal paraholic k -subgroup of G containing a maximal k -split torus S . Put $U = R_\infty(P)$, $Z = Z_\infty(S)$, and $N = N_\infty(S)$. Then, $N = N_k Z$, so G_k is the disjoint union over ${}_k W$ of double cosets $U_k n_w P_k$ ($w \in {}_k W$), where n_w is a representative in N_k of $w \in {}_k W$. More precisely, if $w \in {}_k W$, there exist two k -subgroups U'_w and U''_w such that $U = U'_w \times U''_w$ (product of k -varieties), the map $U'_w \times P \rightarrow U n_w P$ sending (x, y) onto $x n_w y$ is an isomorphism, and

$$(G/P)_k = G_k/P_k = \bigcup_{w \in {}_k W} \pi((U'_w)_k),$$

where π is the projection $G \rightarrow G/P$. This is called a **relative Bruhat decomposition**. If G is k -split, this gives an ordinary Bruhat decomposition (\rightarrow Section K). If θ is a subset of ${}_k \Delta$, let W_θ be the subgroup of ${}_k W$ generated by reflections defined by the α 's in θ . If θ, θ' are subsets of ${}_k \Delta$, then there is a bijection of double cosets

$$(P_\theta)_k \backslash G_k / (P_{\theta'})_k = W_\theta \backslash {}_k W / W_{\theta'}.$$

(Note that all these properties follow from the fact that (P_k, N_k) is a BN-pair in G_k ; \rightarrow Section R.)

If G is k -isotropic, let G^+ be the subgroup of G_k generated by all U_k , where U runs over the unipotent radicals of the minimal paraholic k -

subgroups of G . Kneser and Tits conjectured that when G is semisimple and simply connected and k arbitrary, $G_k = G^+$. Platonov showed that this conjecture fails for some k -forms of SL_n ($n \geq 2$), but it is true in many cases, e.g., when G is k -split or k -quasi-split (J. Tits, *Sém. Bourbaki* 29e, no. 505, 1976677). In this case, if G is simple as an algebraic group and the cardinality of k is ≥ 4 , then $G_k/Z(G_k)$ is simple as an abstract group. The connected semisimple algebraic k -groups over a finite field k are k -split or k -quasi-split (\rightarrow Section 0). J. Tits has given a reduction process for classifying the groups over an arbitrary field k . He defined the index of the k -group (to a certain extent a generalization of Witt's theorem characterizing a quadratic form by means of its index and anisotropic kernel) and gave all possible indices of the k -groups and also a complete list in the case of local field (J. Tits [30, 31]).

R. Buildings and BN-pairs [25]

The origin of the notations of buildings and BN-pairs lies in an attempt to give a systematic procedure for the geometric interpretation of the semisimple Lie groups and, in particular, the exceptional groups. The theory has various applications to the groups of Lie types. To describe this precisely, we must introduce a number of definitions.

A **complex** A is a set with an order relation \subset , read "is a **face** of" or "is contained in" such that for a given element A , the ordered subset $S(A)$ of all faces of A is isomorphic to the set of all subsets of a set. The subset $S(A)$ of A is called a **simplex** in A . A complex has a smallest element, which we denote by 0. For an element A , the number $\text{rk } A$ of minimal nonzero faces of A is called the **rank** of A . Define $\text{rk } A = \sup\{\text{rk } A' \mid A' \in \Delta\}$. A **morphism** $\alpha: \Delta \rightarrow \Delta'$ (where A and A' are complexes) is a mapping of the underlying sets such that for every $A \in \Delta$, $S(A) \cong S(\alpha(A))$. A subcomplex of A is a complex whose underlying set is a subset of A such that the inclusion is a morphism. If $A \in \Delta$, the **star** $\text{St}(A)$ of A is the set of all elements of A containing A . With the order relation induced from A , $\text{St}(A)$ is a complex. If $B \in \text{St}(A)$, the rank of B in $\text{St}(A)$ is called the **codimension** of A in B and is denoted by $\text{codim}_B A$.

A complex A is called a **chamber complex** if every element is contained in a maximal element, which is called a **chamber**, and if given two chambers C, C' , there exists a sequence of chambers $C = C_0, C_1, \dots, C_m = C'$ such that $\text{codim}_{C_{i-1}}(C_{i-1} \cap C_i) = \text{codim}_{C_i}(C_{i-1} \cap C_i) \leq 1$ for all $i = 1, 2, \dots, m$. A chamber complex is called **thick** (resp. **thin**) if every element of

codimension 1 is contained in at least three (resp. exactly two) chambers. An endomorphism φ of a thin chamber complex A is called a **folding** if $\varphi^2 = \text{id}$ and if any chamber on $\varphi(\Delta)$ is the image of exactly two chambers in A by φ . For any folding φ , $\varphi(\Delta)$ is called a **root**. A thin chamber complex is called a **Coxeter complex** if for any pair of **adjacent** chambers C, C' , i.e., chambers such that $\text{codim}_C(C \cap C') = 1$, there is a root containing C and not C' . Let Σ be a Coxeter complex. For any root $\varphi(\Sigma)$ of Σ , there is only one root $\varphi'(\Sigma)$, called the **opposite** of $\varphi(\Sigma)$, such that $\varphi(\Sigma) \cap \varphi'(\Sigma)$ does not contain any chamber and $\varphi(\Sigma) \cup \varphi'(\Sigma) = \Sigma$; there is also an involutive automorphism called the **reflection** associated with φ that transforms $\varphi(\Sigma)$ onto $\varphi'(\Sigma)$. The group $W(C)$ generated by all reflections of Σ is called the **Weyl group** of Σ , which turns out to be a **Coxeter group**, i.e., there exist sets I and $\{m_{ij}\}$, $(i, j) \in I \times I$, where m_{ij} are integers or ∞ and $m_{ii} = 1$, such that the group is presented by the generators $\{r_i\}_{i \in I}$ and the fundamental relations $(r_i r_j)^{m_{ij}} = 1$, $i, j \in I$, $m_{ij} \neq \infty$. A Coxeter complex is called **irreducible** if it is not the join of two nonempty Coxeter subcomplexes.

A **building** is a thick chamber complex A with a system \mathfrak{A} of Coxeter subcomplexes, called the **apartments** of A such that (i) every two simplexes of A belong to an apartment. (ii) If $\Sigma, \Sigma' \in \mathfrak{A}$, there exists an isomorphism of Σ onto Σ' that fixes $\Sigma \cap \Sigma'$ (elementwise). Since the apartments of a building are isomorphic to each other, we can define the Weyl group, rank, and irreducibility of the building to be those of its apartment. A building with finitely many chambers is said to be of **spherical type**. If a building can be realized as a simplicial decomposition of a Euclidean space, then one says it is of **Euclidean type**.

Examples of buildings are provided by **BN-pairs**. A **BN-pair** or **Tits system** in a group G is a system (B, N) consisting of two subgroups of G such that

- (BNO) B and N generate G ;
- (BN1) $B \cap N = H \triangleleft N$; and
- (BN2) The group $W = N/H$ has a generating set R such that for any $r \in R$ and any $w \in W$,
- (BN2') $rBw \subset BwB \cap BrwB$,
- (BN2'') $rBr \neq B$.

The group W is called the **Weyl group** of the BN-pair. For any subset S of R , let W_S denote the subgroup of W generated by S . Set $P_S = B W_S B$. Then P_S is a subgroup of G and the mapping $S \mapsto P_S$ is a bijection of the lattice of all subsets of R onto the lattice of all subgroups of G containing B . A subgroup of G is called **parabolic** if it is conjugate to some P_S . Let A be the set of all left cosets of all subgroups P_S , $S \subset R$, ordered by the opposite of the inclusion relation. Let G operate on A on

the left. Let Σ be the subset $\{nP_S \mid n \in N, S \subset R\}$ of A , and \mathfrak{A} be the set of G -translates of Σ . Then (A, \mathfrak{A}) is a building whose Weyl group is W .

Let G be a reductive algebraic group defined over a field k and P be a minimal parabolic k -subgroup of G containing a maximal k -split torus S and set $N = N_k(S)$. Then (P_k, N_k) is a BN-pair in G_k . Therefore there is a building of spherical type associated with the group G . Conversely, the buildings of rank ≥ 3 and irreducible spherical type (roughly speaking) all turn out to be associated with simple algebraic or classical groups (J. Tits [25]). This result gives a complete and unified description of structures that were discovered previously in certain cases. For example, the building of type A , gives a projective space (E. Abe, T. Tsuzuku), that of type C_n gives a polar space (Veldkamp), and that of a k -form of type E_6 gives a Cayley space (J. Tits). As an application, one can show that a finite building of rank ≥ 3 and irreducible spherical type is isomorphic to the building of an absolutely simple algebraic group over a finite field.

When k is local (i.e., endowed with a complete discrete valuation whose residue field is Perfect), the reductive group defined over k has another BN-pair such that the associated building is Euclidean. This theory was initiated by N. Iwahori and H. Matsumoto [26], who considered split semisimple groups. Later, quasi-split and classical groups were studied by H. Hijikata, and the theory for the general case was given by F. Bruhat and J. Tits [27, 28]. To distinguish from the usual BN-pair structure, the subgroups conjugate to B in this case are called **Iwahori subgroups**, and parabolic subgroups are called **paraholic subgroups**. The Euclidean buildings are the "ultrametric analogs" of the Riemannian symmetric spaces. In other words, in the study of p -adic simple groups, they play a role similar to that of the symmetric spaces in the theory of simple groups.

References

- [1] A. Borel, Linear algebraic groups, Benjamin, 1969.
- [2] J. E. Humphreys, Linear algebraic groups, Springer, 1975.
- [3] M. Demazure and P. Gabriel, Groupes algébriques I, North-Holland, 1970.
- [4] M. Demazure and A. Grothendieck, Schémas en groupes, Séminaire de géométrie algébrique, 1963/1964, Publ. Math. Inst. HES, 1964 (Lecture notes in math. 153, Springer, 1970).
- [5] C. Chevalley, Théorie des groupes de Lie, Groupes algébriques, Théorèmes généraux sur les algèbres de Lie, Actualités Sci. Ind., 1152 and 1226, Hermann, 1968.
- [6] C. Chevalley, Sur certains groupes simples, Tôhoku Math. J., (2) 7 (1955), 14–66.
- [7] C. Chevalley, Classification des groupes de Lie algébriques, Séminaire C. Chevalley, Inst. H. Poincaré, Univ. Paris, 1956/1958.
- [8] M. Rosenlicht, Some basic theorems on algebraic groups, Amer. J. Math., 78 (1956), 401–443.
- [9] M. Rosenlicht, Some rationality questions on algebraic groups, Ann. Mat. Pura Appl., 43 (1957), 25–50.
- [10] A. Weil, On algebraic groups of transformations, Amer. J. Math., 77 (1955), 355–391.
- [11] A. Weil, On algebraic groups and homogeneous spaces, Amer. J. Math., 77 (1955), 493–512.
- [12] J.-P. Serre, Groupes algébriques et corps de classes, Actualités Sci. Ind., Hermann, 1959.
- [13] J. Dieudonné, Introduction to formal groups, Dekker, 1973.
- [14] N. Bourbaki, Groupes et algèbres de Lie, Hermann, 1968, chs. 446.
- [15] J.-P. Serre, Cohomologie galoisienne, Lecture notes in math. 5, Springer, 1964.
- [16] I. Satake, Classification theory of semi-simple algebraic groups, Dekker, 1971.
- [17] J. Tits, Algebraic and abstract simple groups, Ann. Math., (2) 80 (1964), 313–329.
- [18] A. Weil, Algebras with involutions and the classical groups, J. Indian Math. Soc., 24 (1960), 589–623.
- [19] A. Weil, Adeles and algebraic groups, Lecture notes, Institute for Advanced Study, Princeton, 1961.
- [20] A. Borel, Introduction aux groupes arithmétiques, Hermann, 1969.
- [21] A. Borel and Harish-Chandra, Arithmetic subgroups of algebraic groups, Ann. Math., (2) 75 (1962), 485–535.
- [22] G. D. Mostow and T. Tamagawa, On the compactness of arithmetically defined homogeneous spaces, Ann. Math., (2) 76 (1962), 446–463.
- [23] T. Ono, On the Tamagawa number of algebraic tori, Ann. Math., (2) 78 (1963), 44–73.
- [24] A. Borel and J. Tits, Groupes réductifs, Publ. Math. Inst. HES, 27 (1965), 555–611. Complements, Publ. Math. Inst. HES, 41 (1972), 253–276.
- [25] J. Tits, Buildings of spherical type and finite BN-pairs, Lecture notes in math. 386, Springer, 1974.
- [26] N. Iwahori and H. Matsumoto, On some Bruhat decompositions and the structure of p -adic Chevalley groups, Publ. Math. Inst. HES, 25 (1965), 5–48.

[27] H. Hijikata, On the structure of semi-simple algebraic groups over valuation fields I, Japan. J. Math. 1 (1975), 225-300.

[28] F. Bruhat and J. Tits, Groupes réductifs sur un corps local I. Données radicielles valuées, Publ. Math. Inst. HES, 41 (1972), 55-252. II. Schémas en groupes. Existence d'une donnée radicielle valuée, Publ. Math. Inst. HES, 60 (1984), 5-184.

[29] Colloque sur la théorie des groupes algébriques, C. B. R. M., Brussels, 1962.

[30] Algebraic groups and discontinuous subgroups, Amer. Math. Soc. Proc. Symposia in Pure Math., 9, 1966.

[31] Automorphic forms, representations and L-functions, Amer. Math. Soc. Proc. Symposia in Pure Math., 33, 1979.

[32] T. A. Springer, Linear algebraic groups, Birkhäuser, 1981.

[33] G. P. Hochschild, Basic theory of algebraic groups and Lie algebras, Springer, 1981.

14 (V.13) Algebraic Number Fields

A. Introduction

A complex number that satisfies an algebraic equation with rational integral coefficients is said to be an **algebraic number**. If the coefficient of the term of highest degree of the equation is 1, this algebraic number is said to be an **algebraic integer**. The set A of all algebraic numbers is a field which is the \dagger algebraic closure of the rational number field \mathbb{Q} in the complex number field \mathbb{C} . The set I of all algebraic integers is an \dagger integral domain which contains the integral domain \mathbb{Z} of all the rational integers. The \dagger field of quotients of I is A .

B. Principal Order

An extension field k of \mathbb{Q} of finite degree (which we shall always suppose to be contained in \mathbb{C}) is said to be an **algebraic number field** of finite degree, and k is a subfield of A . The intersection $\mathfrak{o} = k \cap I$ is an integral domain whose field of quotients is k ; \mathfrak{o} is called the **principal order** of k . (More generally, a subring R of \mathfrak{o} containing 1 is said to be an **order** of k if the field of quotients of R is k . The set \mathfrak{f} of all elements y of \mathfrak{o} such that $y\mathfrak{o} \subset R$ is an **ideal** of \mathfrak{o} ; in addition \mathfrak{f} is called the **conductor** of R .) Let n be the degree of k over \mathbb{Q} . Then the additive group of the principal order \mathfrak{o} of k is a \dagger free Abelian group of \dagger rank n . A \dagger basis $(\omega_1, \dots, \omega_n)$

of \mathfrak{o} as a free Abelian group (or \mathbb{Z} -module) is said to be a **minimal basis** of \mathfrak{o} (or of k). Let $\omega_j^{(i)}$ ($i = 1, \dots, n$) be \dagger conjugate elements of ω_j over \mathbb{Q} , and let $A = |\omega_j^{(i)}|$ be the determinant whose (i, j) entry is $\omega_j^{(i)}$. Then $D_k = A^2$ is a rational integer that is independent of the choice of a minimal basis of \mathfrak{o} . D_k is called the **discriminant** of k . If $k \neq \mathbb{Q}$, then $|D_k| > 1$ (**Minkowski's theorem**, 1891). For any given rational integer m there are only a finite number of algebraic number fields whose discriminants are equal to m (C. Hermite and H. Minkowski, 1896). The proof of these theorems depends on the methods of geometry of numbers (- 182 Geometry of Numbers).

C. Ideals of the Principal Order

An \dagger ideal \mathfrak{a} of the principal order \mathfrak{o} is said to be an **integral ideal** of k . In particular, a prime ideal ($\neq 0$) of \mathfrak{o} is called simply a prime ideal of k . The domain \mathfrak{o} is not necessarily a \dagger principal ideal ring but is always a \dagger Dedekind domain. That is, every ideal \mathfrak{a} of \mathfrak{o} is uniquely expressed (up to the order of the factors) as a finite product of powers of prime ideals of \mathfrak{o} . This theorem is called the **fundamental theorem of the principal order**.

The quotient ring $\mathfrak{o}/\mathfrak{a}$ of \mathfrak{o} by an ideal \mathfrak{a} ($\neq 0$) of \mathfrak{o} is a finite ring. The number of elements of $\mathfrak{o}/\mathfrak{a}$ is called the **absolute norm** of \mathfrak{a} and is denoted by $N(\mathfrak{a})$. We have $N(\mathfrak{a}\mathfrak{b}) = N(\mathfrak{a})N(\mathfrak{b})$. Every prime ideal \mathfrak{p} ($\neq 0$) of \mathfrak{o} is a \dagger maximal ideal of \mathfrak{o} , and $\mathfrak{o}/\mathfrak{p}$ is a finite field. Let the \dagger characteristic of $\mathfrak{o}/\mathfrak{p}$ be p , where p is a prime number. Then $\mathfrak{o}/\mathfrak{p}$ is a finite extension of the \dagger prime field $\mathbb{Z}/p\mathbb{Z}$. Let the degree of $\mathfrak{o}/\mathfrak{p}$ over $\mathbb{Z}/p\mathbb{Z}$ be f . Then $N(\mathfrak{p}) = p^f$, and f is said to be the **degree** of the prime ideal \mathfrak{p} .

Let s be a complex variable. The (complex-valued) function

$$\zeta_k(s) = \sum_{\mathfrak{a}} 1/N(\mathfrak{a})^s = \prod_{\mathfrak{p}} 1/(1 - (N(\mathfrak{p}))^{-s})$$

of $s \in \mathbb{C}$ is called the **Dedekind zeta function** of k (R. Dedekind, 1871). Here the summation extends over all ideals \mathfrak{a} of \mathfrak{o} , and the product extends over all prime ideals \mathfrak{p} of \mathfrak{o} . This series converges absolutely for $\text{Re } s > 1$, and the function $\zeta_k(s)$ has a single-valued \dagger analytic continuation to a \dagger meromorphic function on the whole complex plane (- 450 Zeta Functions).

D. Units

An algebraic integer ε of k is said to be a **unit** of k if ε^{-1} is also an algebraic integer. Hence ε is a unit of k if and only if the \dagger Principal ideal

(E) is 0. The set E_k of all units of k forms an Abelian group under multiplication, which is called the **unit group** of k . The set of all elements of E_k of finite order coincides with the set of all the roots of unity contained in k and forms a cyclic group of a finite order w . Let n be the degree of k over Q . Then for each element $\alpha \in k$ there are n conjugate elements $\alpha^{(i)}$ over Q . Let $\alpha^{(i)}$ ($i = 1, \dots, r_1$) be real for any $\alpha \in k$, and let $\alpha^{(r_1+j)}$ and $\alpha^{(r_1+r_2+j)}$ ($j = 1, \dots, r_2$) be pairs of complex conjugates for any $\alpha \in k$. Then we have $n = r_1 + 2r_2$. The unit group E_k of k is the direct product of a cyclic group of order w and the free Abelian multiplicative group of rank $r = r_1 + r_2 = 1$. This theorem is called **Dirichlet's unit theorem** (1846). A basis $(\varepsilon_1, \dots, \varepsilon_r)$ of this free group is called a system of **fundamental units** of k .

Let $l^{(i)}\alpha = \log|\alpha^{(i)}|$ ($i = 1, \dots, r_1$), $l^{(j)}\alpha = 2\log|\alpha^{(j)}|$ ($j = r_1 + 1, \dots, r_1 + r_2$) for $\alpha \in k$. For r elements η_1, \dots, η_r of E_k ,

$$R[\eta_1, \dots, \eta_r] = \begin{vmatrix} l^{(1)}\eta_1 & l^{(1)}\eta_2 & \dots & l^{(1)}\eta_r \\ l^{(2)}\eta_1 & l^{(2)}\eta_2 & & l^{(2)}\eta_r \\ \vdots & \vdots & \ddots & \vdots \\ l^{(r)}\eta_1 & l^{(r)}\eta_2 & & l^{(r)}\eta_r \end{vmatrix}$$

is called the **regulator** of (η_1, \dots, η_r) (Dedekind). In order for η_1, \dots, η_r to be multiplicatively independent, it is necessary and sufficient that $R[\eta_1, \dots, \eta_r] \neq 0$. The absolute value of $R[\eta_1, \dots, \eta_r]$ takes the minimum positive value R for fundamental units $(\varepsilon_1, \dots, \varepsilon_r)$. $R = R[\varepsilon_1, \dots, \varepsilon_r]$ is independent of the choice of fundamental units $(\varepsilon_1, \dots, \varepsilon_r)$ of k . R is called the **regulator** of k . In general, $R[\eta_1, \dots, \eta_r]/R$ is equal to the index $[E_k : H]$ of the group H generated by the roots of unity in k and η_1, \dots, η_r . H. W. Leopoldt conjectured that units in k , which are multiplicatively independent over Z , remain multiplicatively independent over Z_p (the ring of p -adic integers) when they are considered as elements of the tensor product $k \otimes Q_p$ over Q . This conjecture was affirmatively proved in some special cases by J. Ax (Illinois J. Math., 9 (1965)) and others.

If k/Q is a Galois extension, there exists a unit ε of k such that the conjugates of ε over Q contain r multiplicatively independent units (Minkowski's theorem).

E. Ideal Classes

An \mathfrak{o} -module contained in k (i.e., $\mathfrak{a}\mathfrak{o} \subset \mathfrak{a}$) such that $\alpha\mathfrak{a} \subset \mathfrak{o}$ holds for some element $\alpha (\neq 0)$ of k is said to be a **fractional ideal** of k . For two fractional ideals $\mathfrak{a}, \mathfrak{b}$ of k the "product" $\mathfrak{a}\mathfrak{b}$ defined by $\{\sum \alpha_i \beta_i \text{ (finite sum)} \mid \alpha_i \in \mathfrak{a}, \beta_i \in \mathfrak{b}\}$ is also a fractional ideal. Thus the set of the

fractional ideals of k forms a multiplicative commutative semigroup. For a fractional ideal \mathfrak{a} the set $\mathfrak{a}^{-1} = \{\alpha \in k \mid \alpha\mathfrak{a} \subset \mathfrak{o}\}$ is also a fractional ideal of k , and we have $\mathfrak{a}\mathfrak{a}^{-1} = \mathfrak{o}$. Thus the set of all nonzero fractional ideals of k forms an Abelian group \mathfrak{I}_k under multiplication with \mathfrak{o} as identity. Each fractional ideal $\mathfrak{a} (\neq \mathfrak{o})$ is uniquely expressed as a finite product of powers of prime ideals, if we admit negative powers. Namely, \mathfrak{I}_k is a free Abelian multiplicative group with the set of all prime ideals as basis. Given fractional ideals \mathfrak{a} and \mathfrak{b} , we say that \mathfrak{a} is **divisible** by \mathfrak{b} if $\mathfrak{a} \subset \mathfrak{b}$; in this case, we call \mathfrak{b} a **divisor** of \mathfrak{a} and \mathfrak{a} a **multiple** of \mathfrak{b} . Also, $\mathfrak{a} \subset \mathfrak{b}$ if and only if there exists an integral ideal \mathfrak{c} such that $\mathfrak{a} = \mathfrak{b}\mathfrak{c}$. Given fractional ideals $\mathfrak{a} = \prod \mathfrak{p}_i^{e_i}$ and $\mathfrak{b} = \prod \mathfrak{q}_j^{f_j}$ ($e_i \neq 0, f_j \neq 0$), we say that \mathfrak{a} and \mathfrak{b} are **relatively prime** if $\{\mathfrak{p}_i\}$ and $\{\mathfrak{q}_j\}$ are disjoint. Usually a fractional ideal of k is simply called an ideal of k .

For an element $\alpha (\neq 0)$ of k , $(\alpha) = \alpha\mathfrak{o}$ is a (fractional) ideal of k , and (α) is said to be a **principal ideal** of k . The set P_k of all principal ideals (α) ($\alpha \in k, \alpha \neq 0$) is a subgroup of \mathfrak{I}_k . Since $(\alpha) = \mathfrak{o}$ is equivalent to $\alpha \in E_k$, we have $P_k \cong k^*/E_k$, where k^* is the multiplicative group of all nonzero elements of k .

Each coset of \mathfrak{I}_k modulo P_k is called an **ideal class** of k , and the group $\mathfrak{C}_k = \mathfrak{I}_k/P_k$ is called the **ideal class group** of k . Each ideal class contains an integral ideal \mathfrak{a} with $N(\mathfrak{a}) \leq \sqrt{|D_k|}$ (more precisely, with $N(\mathfrak{a}) \leq (4/\pi)^{r_2} (n!/n^n) \sqrt{|D_k|}$). From this it follows that \mathfrak{C}_k is a finite Abelian group. The order h of \mathfrak{C}_k is called the **class number** of k . For the calculation of the class number the \dagger residue at the pole $s = 1$ of the Dedekind zeta function is used. Namely,

$$\lim_{s \rightarrow 1+0} (s-1)\zeta_k(s) = gh,$$

$$g = 2^{r_1+r_2} \pi^{r_2} R_k / w_k \sqrt{|D_k|},$$

where R_k is the regulator of k and w_k is the number of roots of unity in k (Dedekind, 1877). This formula is used, in particular, for the computation of the class numbers of \dagger quadratic fields and \dagger cyclotomic fields (- 347 Quadratic Fields). The class numbers of cubic and quartic (real) \dagger cyclic fields over Q were computed by H. Hasse in the case where the \dagger conductor of k/Q is less than 100 (*Abh. Deutsch. Akad. Wiss. Berlin*, 2 (1948)). Hasse has also given a detailed computation of the class number of cyclotomic fields [15]. In general, let the degree $n = [k: Q]$ be fixed and let $|D_k| \rightarrow \infty$. Then

$$\lim (\log(h_k R_k) / \log \sqrt{|D_k|}) = 1.$$

(This formula was proved for $n = 2$ by C. L. Siegel, 1935, and for general n by R. Brauer, *Amer. J. Math.*, 69 (1947).)

F. Valuations

All the \dagger Archimedean and \dagger non-Archimedean valuations of an algebraic number field k can be obtained as follows (- 439 Valuations):

Archimedean Valuations. Let $\eta = [k: \mathbf{Q}]$, and let η conjugates of $\alpha \in k$ be $\alpha^{(1)}, \dots, \alpha^{(\eta)}$ such that $\alpha^{(i)}$ ($i = 1, \dots, r_1$) is real, and $\alpha^{(r_1+j)}$ and $\alpha^{(r_1+r_2+j)}$ ($j = 1, \dots, r_2$) are pairs of complex conjugates. We write $|\alpha|_i = |\alpha^{(i)}|$ ($j = 1, r_1 + r_2$); these are Archimedean valuations of k that are not mutually equivalent. The equivalence classes of these valuations are denoted by $p_{\infty}^{(1)}, \dots, p_{\infty}^{(r_1+r_2)}$, respectively, and are called the \dagger infinite prime divisors of k . The first r_1 infinite prime divisors are called \dagger real and the remaining r_2 are called \dagger imaginary (or complex). The valuations of k defined by

$$|\alpha|_{p_j} = |\alpha|_j, \quad j = 1, \dots, r_1, \\ = |\alpha|_j^2, \quad j = r_1 + 1, \dots, r_1 + r_2,$$

are called \dagger normal valuations of k . Here $|_{p_j}$ ($j = r_1 + 1, \dots, r_1 + r_2$) are valuations in the wider sense. If $r_1 = \eta$ we call k a **totally real field**, and if $r_1 = 0$ we call k a **totally imaginary field**.

Non-Archimedean Valuations. Let p be a prime ideal of k and α an element of k . Let $(\alpha) = p^a b$, where p and b are relatively prime. Put $v_p(\alpha) = a$. Then for any constant ρ ($0 < \rho < 1$),

$$|\alpha|_p = \rho^{v_p(\alpha)}$$

is a non-Archimedean valuation of k . This valuation of k is called the \dagger p -adic valuation of k ; p -adic valuations for different prime ideals are mutually inequivalent. The valuation $|\alpha|_p$ with $\rho = (N(p))^{-1}$ is called a \dagger normal valuation of k . The equivalence class of valuations containing $|_p$ is denoted by the same letter p and is called a \dagger finite prime divisor of k .

A formal finite product of powers of finite or infinite prime divisors $m^* = \prod p_i^{e_i}$ is called a **divisor** of k . If all $e_i \geq 0$, then m^* is called an **integral divisor** of k . Given divisors $m^* = \prod p_i^{e_i}$ and $n^* = \prod p_i^{f_i}$, we write $m^* \cdot n^*$ if $e_i \leq f_i$ ($i = 1, 2, \dots$).

Any valuation of k is equivalent to one of the valuations defined previously (A. Ostrowski, 1918; E. Artin, 1932). For any element α ($\neq 0$) of k the \dagger product formula $\prod_p |\alpha|_p = 1$ holds, where p runs over all finite and infinite prime divisors of k and $|_p$ are the normal valuations of k . Conversely, let k be a field, and let $V = \{ |_p \}$ be a set of inequivalent valuations of k such that (i) for any $a \in k$ ($a \neq 0$) $|\alpha|_p \neq 1$ holds only for a finite number of p in V ; (ii) the product formula $\prod_p |\alpha|_p = 1$ ($\alpha \in k$,

$\alpha \neq 0$) holds; and (iii) there is at least one Archimedean valuation in V . Then k is an algebraic number field and V is the set of all the prime divisors of k (Artin and G. Whaples, *Bull. Amer. Math. Soc.*, 51 (1945)).

G. Ideal Classes in the Narrow Sense

For $a, \beta \in k$, the expression $a \equiv \beta \pmod{p_{\infty}^{(i)}}$ means $\alpha^{(i)} \beta^{(i)} > 0$ for a real infinite prime divisor $p_{\infty}^{(i)}$ and $\alpha^{(i)} \beta^{(i)} \neq 0$ for an imaginary infinite prime divisor $p_{\infty}^{(i)}$. We call an element $\alpha \in k$ **totally positive** if all real conjugates $\alpha^{(i)}$ ($i = 1, \dots, r_1$) are positive. In the notation just given, this means $a \equiv 1 \pmod{p_{\infty}^{(i)}}$ ($i = 1, \dots, r_1$). The set of all principal ideals (a) generated by totally positive elements $a \in k$ is a multiplicative subgroup P_k^+ of P_k . Each coset of \mathfrak{J}_k modulo P_k^+ is called an **ideal class of k in the narrow sense**. Let E_k^+ be the group of all totally positive units of k . Then we have $(\mathfrak{J}_k : P_k^+) = h 2^{r_1} / (E_k : E_k^+)$.

H. Multiplicative Congruence

Let m be an integral ideal of k , and let $k^*(m)$ be the multiplicative group of all elements a in k such that (a) is relatively prime to m . Any element $\alpha \in k^*(m)$ can be expressed in the form $\alpha = \beta/\gamma$ such that $\beta, \gamma \in o$ and $(\beta), (\gamma)$ are relatively prime to m .

Consider an integral divisor $m^* = m \prod p_{\infty}^{(i)}$ which is a formal product of m and infinite prime divisors $p_{\infty}^{(i)}$ of k . We call m the finite part of m^* . Given an element $\alpha \in k^*(m)$ and elements $\beta, \gamma \in k^*(m) \cap o$ such that $a = \beta/\gamma$, we set $a \equiv 1 \pmod{m^* n^*}$ if $\beta \equiv \gamma \pmod{m n}$, and $a \equiv 1 \pmod{p_{\infty}^{(i)}}$. The set of all a in $k^*(m)$ such that $a \equiv 1 \pmod{m^*}$ forms a multiplicative group. We write $a \equiv \beta \pmod{m^*}$ for $a, \beta \in k$ if $\alpha/\beta \in k^*(m)$ and $\alpha/\beta \equiv 1 \pmod{m^*}$. This congruence is called the **multiplicative congruence**. In the following discussion we shall write $\text{mod } m^*$ for $\text{mod } m^*$.

We denote by $\mathfrak{J}_k(m)$ the group of all ideals of k that are relatively prime to an integral ideal m , and by $S(m^*)$ the group of all principal ideals (a) such that $a \in k^*(m)$, $a \equiv 1 \pmod{m^*}$; $S(m^*)$ is known as the **ray** modulo m^* . Any subgroup H of $\mathfrak{J}_k(m)$ which contains $S(m^*)$ is called an **ideal group** modulo m^* , and the factor group $\mathfrak{J}_k(m)/H$ is called a **group of congruence classes** of ideals modulo m^* .

If $n^* \mid m^*$ for integral divisors m^* and n^* of k , then $\mathfrak{J}_k(m) \subset \mathfrak{J}_k(n)$ and $S(m^*) \subset S(n^*)$. If H is an ideal group modulo n^* , then $Q(H) = H \cap \mathfrak{J}_k(m)$ is an ideal group modulo m^* , and we have $\mathfrak{J}_k(n)/H \cong s.(m)/@.(H)$. For any given ideal group H_0 modulo m^* there is a smallest

integral divisor \mathfrak{f}^* such that $\mathfrak{f}^* \mid \mathfrak{m}^*$, and there exists an ideal group H modulo \mathfrak{f}^* with $\Phi(H) = H_0$ (i.e., if there is an ideal group H' modulo \mathfrak{n}^* with $\Phi(H') = H_0$, then $\mathfrak{f}^* \mid \mathfrak{n}^*$). We call \mathfrak{f}^* the **conductor** of the ideal group H . The notion of multiplicative congruence is used in \dagger class field theory and in the theory of \dagger norm-residue symbols.

I. Ideal Theory for Relative Extensions

If an algebraic number field K has a subfield k , we say that K/k is a **relative algebraic number field**. Let \mathfrak{D} be the principal order of K . For a (fractional) ideal \mathfrak{a} of k , $\mathfrak{D}\mathfrak{a}$ is an ideal of K . We write $\mathfrak{D}\mathfrak{a} = E(\mathfrak{a})$ and call $E(\mathfrak{a})$ the **extension** of \mathfrak{a} to K . For ideals $\mathfrak{a}, \mathfrak{b}$ of k , we have $E(\mathfrak{a}\mathfrak{b}) = E(\mathfrak{a})E(\mathfrak{b})$ and $E(\mathfrak{a}) \cap k = \mathfrak{a}$.

Let $\Psi_i: K \rightarrow \mathbb{C}$ be k -isomorphisms ($i = 1, \dots, n$), where $n = [K:k]$. We write $K^{(i)} = \Psi_i(K)$ and $A^{(i)} = Y_i(A)$ for $A \in K$. For an ideal \mathfrak{A} of K , $\mathfrak{A}^{(i)} = \{A^{(i)} \mid A \in \mathfrak{A}\}$ is an ideal of $K^{(i)}$, and $\mathfrak{A}^{(i)}$ is called the **conjugate ideal** of \mathfrak{A} in $K^{(i)}$. Let L be the composite field of $K^{(1)}, \dots, K^{(n)}$. Then the ideal generated by $\mathfrak{A}^{(1)} \dots \mathfrak{A}^{(n)}$ in L is the extension of an ideal \mathfrak{a} of k . We write $\mathfrak{a} = N_{K/k}(\mathfrak{A})$ and call \mathfrak{a} the **relative norm** of \mathfrak{A} over k . We have $N_{K/k}(\mathfrak{A}\mathfrak{B}) = N_{K/k}(\mathfrak{A})N_{K/k}(\mathfrak{B})$ and $N_{K/k}(E(\mathfrak{a})) = \mathfrak{a}$ (for an ideal \mathfrak{a} of k). In particular, for $k = \mathbb{Q}$, $N_{K/\mathbb{Q}}(\mathfrak{A}) = (N(\mathfrak{A}))$.

Let \mathfrak{p} be a prime ideal of k . Then $E(\mathfrak{p}) = \mathfrak{P}_1^{e_1} \mathfrak{P}_2^{e_2} \dots \mathfrak{P}_g^{e_g}$ in \mathfrak{D} , where $\mathfrak{P}_1, \dots, \mathfrak{P}_g$ are prime ideals of K . Let f_i be the degree of the finite field $\mathfrak{D}/\mathfrak{P}_i$ over $\mathfrak{o}/\mathfrak{p}$. Then $N_{K/k}(\mathfrak{P}_i) = \mathfrak{p}^{f_i}$; f_i is called the **relative degree** of \mathfrak{P}_i over k , and e_i is called the (relative) **ramification index** of \mathfrak{P}_i over k . We have the relation $n = \sum_{i=1}^g e_i f_i$ between these numbers. If $e_i = e_g = 1$, the prime ideal \mathfrak{p} is said to be **unramified** for K/k . Otherwise, \mathfrak{p} is said to be **ramified** for K/k . If every prime ideal of k is unramified for K/k , we call K/k an **unramified extension**. (For an infinite prime divisor \mathfrak{p}_∞ of k we write $\mathfrak{p}_\infty = \prod_{i=1}^g \mathfrak{P}_\infty^{(i)e_i}$ if the Archimedean valuation $|\cdot|_{\mathfrak{p}_\infty}$ of k can be extended to g Archimedean valuations $|\cdot|_{\mathfrak{P}_\infty^{(i)}} (i = 1, \dots, g)$ of K , where $e_i = 2$ if $\mathfrak{P}_\infty^{(i)}$ is imaginary and \mathfrak{p}_∞ is real, $e_i = 1$ otherwise.)

J. Relative Differents and Relative Discriminants

Let K/k be a relative algebraic number field and $\mathfrak{o}, \mathfrak{o}_K$ be the principal orders of k, K , respectively. Put $\mathfrak{M} = \{A \in K \mid \text{Tr}_{K/k}(A\mathfrak{D}) \subset \mathfrak{o}\}$, where $\text{Tr}_{K/k}$ is the \dagger trace (- 149 Fields J). Then \mathfrak{M} is a (fractional) ideal of K and $\mathfrak{M}^{-1} = \mathfrak{D}_{K/k}$ is an integral ideal of K ; $\mathfrak{D}_{K/k}$ is called the **rela-**

tive different of K over k . When $k = \mathbb{Q}$, $\mathfrak{D}_{K/\mathbb{Q}}$ is simply called the **different** of K . For $L \supset K \supset k$, we have the **chain theorem**: $\mathfrak{D}_{L/k} = \mathfrak{D}_{L/K} \mathfrak{D}_{K/k}$.

Let the conjugates of $A \in K$ over k be $A^{(1)}, \dots, A^{(n)}$, and assume that $A^{(1)} = A$. Put $\delta_{K/k}(A) = \prod_{i=2}^n (A - A^{(i)})$ for $A \in K$. If $A \in \mathfrak{D}$, then $\delta_{K/k}(A) \in \mathfrak{D}_{K/k}$. $\mathfrak{D}_{K/k}$ is generated by $\{\delta_{K/k}(A) \mid A \in \mathfrak{D}\}$. The integral ideal $\mathfrak{E}^{(i)}$ generated by $\{A - A^{(i)} \mid A \in \mathfrak{D}\}$ in the field $L = K^{(1)}K^{(2)} \dots K^{(n)}$ was called an *element* by D. Hilbert. We also have $\mathfrak{D}_{K/k} = \mathfrak{E}^{(2)}\mathfrak{E}^{(3)} \dots \mathfrak{E}^{(n)}$. The integral ideal $\mathfrak{d}_{K/k} = N_{K/k}(\mathfrak{D}_{K/k})$ of k is called the **relative discriminant** of K/k . If $k = \mathbb{Q}$, $\mathfrak{d}_{K/\mathbb{Q}} = (D_K)$.

For the relative different $\mathfrak{D}_{K/k}$ to be divisible by a prime ideal \mathfrak{P} of K , it is necessary and sufficient that $E(\mathfrak{p}) = \mathfrak{P}^e \mathfrak{P}_2^{e_2} \dots \mathfrak{P}_g^{e_g}$ with $e > 1$, where $\mathfrak{p} = \mathfrak{P} \cap k$ (**Dedekind's discriminant theorem**, 1882). Hence a prime ideal \mathfrak{p} of k is ramified for K/k if and only if \mathfrak{p} divides the relative discriminant $\mathfrak{d}_{K/k}$; there are thus only a finite number of prime ideals of k which ramify for K/k . In particular, K/k is unramified if and only if $\mathfrak{d}_{K/k} = 0$.

K. Arithmetic of Galois Extensions

Let K/k be a relative algebraic number field such that K is a \dagger Galois extension of k of degree n , and let G be the \dagger Galois group of K/k . Let $\mathfrak{o}, \mathfrak{D}$ be the principal order of k, K , respectively. The conjugate ideals of an ideal \mathfrak{A} of K are given by $\mathfrak{A}^\sigma = \{A^\sigma \mid A \in \mathfrak{A}\} (\sigma \in G)$. If $N_{K/k}(\mathfrak{A}) = \mathfrak{a}$, then $E(\mathfrak{a}) = \prod_{\sigma \in G} \mathfrak{A}^\sigma$. For a prime ideal \mathfrak{p} of k , $E(\mathfrak{p}) = (\mathfrak{P}_1 \mathfrak{P}_2 \dots \mathfrak{P}_g)^e$, where $N_{K/k}(\mathfrak{P}_i) = \mathfrak{p}^{f_i} (i = 1, \dots, g)$, $n = efg$, and $\mathfrak{P}_1, \dots, \mathfrak{P}_g$ are mutually conjugate prime ideals of K over k .

Hilbert (1894) developed the decomposition theory of a prime ideal \mathfrak{p} of k for a Galois extension K/k in terms of the Galois group G as follows: Let \mathfrak{P} be a prime ideal of \mathfrak{D} . Then

$$Z = \{\sigma \in G \mid \mathfrak{P}^\sigma = \mathfrak{P}\}$$

is a subgroup of the Galois group G of K/k . Z is called the **decomposition group** of \mathfrak{P} over k . Let $G = \bigcup_i Z\tau_i$ be the left coset decomposition of G . Then $\mathfrak{P}_i = \mathfrak{P}^{\tau_i} (i = 1, \dots, g)$ are all the conjugate ideals of \mathfrak{P} over k .

The subgroup

$$T = \{\sigma \in Z \mid A^\sigma \equiv A \pmod{\mathfrak{P}}, A \in \mathfrak{D}\}$$

of the decomposition group Z is normal, and T is called the **inertia group** of \mathfrak{P} over k . The quotient group Z/T is a cyclic group of order f (the relative degree of \mathfrak{P}). There exists an element σ of Z such that

$$A^\sigma \equiv A^{N(\mathfrak{p})} \pmod{\mathfrak{P}}, \quad A \in \mathfrak{D},$$

and σ is uniquely determined mod T ; σT gen-

erates the cyclic group Z/T . This σ is called the **Frobenius substitution** (or **Frobenius automorphism**) of \mathfrak{P} over k . For $m = 1, 2, \dots$,

$$V^{(m)} = \{\sigma \in Z \mid A^\sigma \equiv A \pmod{\mathfrak{P}^{m+1}}, A \in \mathfrak{D}\}$$

are normal subgroups of Z ; the group $V^{(m)}$ is called the m th **ramification group** of \mathfrak{P} over k . Let

$$\begin{aligned} V^{(0)} &= \dots = V^{(v_1)} \supsetneq V^{(v_1+1)} = \dots \\ &= V^{(v_2)} \supsetneq \dots \supsetneq V^{(v_{r-1}+1)} = \dots \\ &= V^{(v_r+1)} \supsetneq V^{(v_r+1)} = 1 \end{aligned}$$

Let $V_\rho = V^{(v_\rho+1)}$ ($\rho = 0, 1, \dots, r$), where $v_0 = -1$. In particular, $V_0 = V^{(0)} = T$. The integers v_1, v_2, \dots are called the **ramification numbers** of \mathfrak{P} . The group T/V_1 is isomorphic to a subgroup of the multiplicative group of the finite field $\mathfrak{D}/\mathfrak{P}$. Hence T/V_1 is a cyclic group whose order e_0 is a divisor of $N(\mathfrak{P}) - 1$. The group V_m/V_{m+1} ($m \geq 1$) is isomorphic to a subgroup of the additive group of the finite field $\mathfrak{D}/\mathfrak{P}$. Hence V_m/V_{m+1} is an Abelian group of type (p, p, \dots, p) whose order divides $N(\mathfrak{P})$. From $e = |T| = (T: V_1)/|V_1|$ it follows that $e = e_0 p^h$, $(e_0, p) = 1$. Here G_1 denotes the order of a finite group G . Hence the decomposition group of \mathfrak{P} is a \dagger solvable group. The relation between the ramification numbers for K/k and those for an intermediate Galois extension F/k was completely determined by J. Herbrand (*J. Math. Pures Appl.*, 10 (1931)) [III].

Let \mathfrak{P}^d be the \mathfrak{P} -component of the relative different $\mathfrak{D}_{K/k}$ of a Galois extension K/k . Then

$$d = \sum_{\rho=0}^{r-1} (v_{\rho+1} - v_\rho)(|V_\rho| - 1) = \sum_{i=0}^{v_r} (|V^{(i)}| - 1).$$

In particular, $d = 0$ if $T = 1$, and $d = e - 1$ if $V^{(1)} = 1$.

Let k_Z, k_T , and $k_{V^{(m)}}$ be the intermediate fields which correspond to the subgroups Z, T , and $V^{(m)}$, respectively, in the sense of \dagger Galois theory; the fields k_Z, k_T , and $k_{V^{(m)}}$ are called the **decomposition field**, the **inertia field**, and the m th **ramification field** of \mathfrak{P} , respectively. Let \mathfrak{P} be a prime ideal of K containing \mathfrak{p} , and let \mathfrak{p}_Z and \mathfrak{p}_T be prime ideals in k_Z and k_T such that $\mathfrak{p}_Z = \mathfrak{P} \cap k_Z$ and $\mathfrak{p}_T = \mathfrak{P} \cap k_T$. Then we have $E(\mathfrak{p}) = \mathfrak{p}_Z \mathfrak{p}_Z^{(2)} \mathfrak{p}_Z^{(g)}$ for k_Z/k ; $E(\mathfrak{p}_Z) = \mathfrak{p}_T$ and $N_{k_T/k_Z}(\mathfrak{p}_T) = \mathfrak{p}_Z^f$ for k_T/k_Z ; and $E(\mathfrak{p}_T) = \mathfrak{P}^e$ for K/k_T .

If a prime ideal \mathfrak{p} of k is unramified for a Galois extension K/k then we have $E(\mathfrak{p}) = \mathfrak{P}_1 \mathfrak{P}_2 \mathfrak{P}_g$, $\mathfrak{P}_i = \mathfrak{P}^{f_i}$ ($i = 1, \dots, g$), $N_{N/k}(\mathfrak{P}_i) = \mathfrak{p}^{f_i}$, and $n = fg$. The Frobenius automorphism $\sigma_i: A^{\sigma_i} \equiv A^{N^{(v)}} \pmod{\mathfrak{P}_i}$ ($A \in \mathfrak{D}$) for the prime ideal \mathfrak{P}_i is uniquely determined, and its order is f_i . Since $\mathfrak{P}_i = \mathfrak{P}^{f_i}$, we have $\sigma_i = \tau_i^{-1} \sigma_1 \tau_i$. Hence $\sigma_1, \dots, \sigma_g$ belong to the same \dagger conjugate class of G . In particular, if G is an Abelian group, then $\sigma_1 = \dots = \sigma_g$ and

$$A^{\sigma_1} \equiv A^{N^{(v)}} \pmod{\mathfrak{p}}, \quad A \in \mathfrak{D}.$$

We then write

$$\sigma_1 = \left(\frac{k}{\mathfrak{p}} \right) \in G$$

and call this symbol the **Artin symbol** for \mathfrak{p} for the Abelian extension K/k . For an ideal $\mathfrak{a} = \prod \mathfrak{p}^{e_i}$ of k that is relatively prime to the relative discriminant of K/k , we define

$$\left(\frac{K/k}{\mathfrak{a}} \right) = \prod \left(\frac{K/k}{\mathfrak{p}} \right)^{e_i} \quad (\in G)$$

Evidently, we have

$$\left(\frac{K/k}{\mathfrak{ab}} \right) = \left(\frac{K/k}{\mathfrak{a}} \right) \left(\frac{K/k}{\mathfrak{b}} \right).$$

The arithmetic of quadratic fields (\rightarrow 347 Quadratic Fields) and the arithmetic of cyclotomic fields (\rightarrow Section L) have been developed since the 19th Century.

L. Arithmetic of Cyclotomic Fields

A complex number ζ whose m th power is 1 but whose m' th power is not 1 for $m' < m$ is called an m th **primitive root of unity**. There are $\varphi(m)$ primitive roots of unity: $\exp(2\pi i r/m)$ ($(r, m) = 1$), where φ is \dagger Euler's function. These $\varphi(m)$ primitive roots of unity are the zeros of an irreducible polynomial over \mathbb{Q} of degree $\varphi(m)$:

$$F_m(X) = \prod_{d|m} (X^{m/d} - 1)^{\mu(d)},$$

where μ is the \dagger Möbius function. The coefficient of the highest term of $F_m(X)$ is 1, and the other coefficients are all rational integers.

$F_m(X)$ is called a **cyclotomic polynomial**. An example is

$$\begin{aligned} F_{12}(X) &= (X^{12} - 1)(X^2 - 1)/(X^6 - 1)(X^4 - 1) \\ &= X^4 - X^2 + 1. \end{aligned}$$

The algebraic number field $K_m = \mathbb{Q}(\zeta_m)$ ($\zeta_m = \exp(2\pi i/m)$) obtained by adjoining an m th primitive root of unity to \mathbb{Q} is a Galois extension over \mathbb{Q} of degree $\varphi(m)$ whose Galois group G is isomorphic to the multiplicative Abelian group of \dagger reduced residue classes of \mathbb{Z} modulo m : $G = \{\sigma_r \mid \zeta_m^{\sigma_r} = \zeta_m^r, (r, m) = 1\}$. K_m is called the m th **cyclotomic field**. Cyclotomic fields are \dagger Abelian extensions of \mathbb{Q} . Conversely, every Abelian extension of \mathbb{Q} is a subfield of a cyclotomic field (**Kronecker's theorem**, 18.53, 1877).

We can choose $(1, \zeta_m, \zeta_m^2, \dots, \zeta_m^{\varphi(m)-1})$ as a minimal basis of K_m . Let $m = l_1^{h_1} l_2^{h_2} \dots l_t^{h_t}$ be the decompositions of m in powers of prime numbers l_1, \dots, l_t . Put $K^{(i)} = K_{l_i^{h_i}}$. Then K_m is the composite field $K_m = K^{(1)} K^{(2)} \dots K^{(t)}$. The different of K_m is given by $\mathfrak{D}_{K_m/\mathbb{Q}} =$

$\mathfrak{D}_{K^{(1)}/Q} \mathfrak{D}_{K^{(2)}/Q} \dots \mathfrak{D}_{K^{(l)}/Q}$, and the discriminant of K_m is given by $D_{K_m} = D_{K^{(1)}}^{n_1} D_{K^{(l)}}^{n_l}$ ($n_i = \varphi(m)/\varphi(l^{h_i})$). If $m = l^h$, the discriminant of K_m is $D_{K_m} = \varepsilon l^a$ ($a = l^{h-1}(hl - h - 1)$), where $\varepsilon = -1$ if $l^h = 4$ or $l \equiv 3 \pmod{4}$, and $\varepsilon = 1$ otherwise. Hence the discriminant of K_m is $D_{K_m} = (\sqrt{-1} m / \prod_{p|m} p^{1/(p-1)})^{\varphi(m)}$. Suppose that $m \not\equiv 2 \pmod{4}$. Then a prime number p ramifies for K_m/Q if and only if p divides m . In particular, if $m = l^h$ ($h > 2$), then only 1 ramifies for K_m/Q and $(l) = [l^{\varphi(m)}]$, $N(l) = l$ (l is explicitly given by $l = (1-i, \dots)$). For $l \neq 2$ the ramification numbers for l are $v_i = l^i$ ($i = 1, 2, \dots$), and the ramification fields are K_l, K_{l^2}, \dots . For $l = 2$ the ramification numbers are 1, 3, 7, and the ramification fields are Q, K_4, K_8 .

In K_m/Q a prime number p ($p \nmid m$) is decomposed as $(p) = \mathfrak{p}_1 \mathfrak{p}_g$, $N(\mathfrak{p}_i) = p^f$ ($i = 1, \dots, g$) and $fg = \varphi(m)$. Here the degree f of \mathfrak{p}_i is determined as the minimal positive integer f such that $p^f \equiv 1 \pmod{m}$. Hence the decomposition law of a prime number p in K_m/Q is determined by its residue class modulo m . This is a prototype form of class field theory (- 59 Class Field Theory).

The class number of the cyclotomic field K_m can be calculated by Dedekind's formula (- Section E; see also Hilbert [4]). Here we shall give the result for $m = l$ (a prime number). Let ζ be a primitive root modulo l . For $\zeta = \exp(2\pi i/l)$, we put

$$\varepsilon = \varepsilon(\zeta) = \left(\frac{1 - \zeta^r}{1 - \zeta} \frac{1 - \zeta^{-r}}{1 - \zeta^{-1}} \right)^{1/2}.$$

Then ε is a unit in K_l . Define an element σ of the Galois group of K_l/Q by $\zeta^\sigma = \zeta^r$, and put $\varepsilon_i = \varepsilon^{\sigma^i}$ ($i = 0, 1, \dots$). Then $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{\rho-1}$ ($\rho = (l-3)/2$) are multiplicatively independent units. That is, the regulator $R[\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{\rho-1}] = E \neq 0$. The units $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{\rho-1}$ are called **circular units**. The class number h of K_l is the product of two factors, $h = h_1 h_2$. Here h_2 is the class number of the real subfield $K'_l = Q(\zeta + \zeta^{-1})$. E. E. Kummer called k_1 the **first factor** and h_2 the **second factor** of the class number h . Let $\chi_1, \chi_2, \dots, \chi_{l-1}$ be the multiplicative characters of the reduced residue classes of \mathbf{Z} modulo l , and let χ_i ($i = 1, \dots, \rho+1$) be the characters among them such that $\chi_i(-1) = -1$. Then

$$h_1 = \frac{(-1)^{\rho+1}}{(2l)^\rho} \prod_{i=1}^{\rho+1} \left(\sum_{a=1}^{l-1} a \chi_i(a) \right), \quad h_2 = \frac{|E|}{R_0}$$

(Kummer, 1850). Here R_0 is the regulator of K'_l . Since circular units belong to K'_l , the class number h_2 of K'_l is equal to the index of the subgroup generated by $\pm 1, \varepsilon_0, \dots, \varepsilon_{\rho-1}$ in the group of units of K'_l . The class number h of K_l is equal to 1 for $l \leq 19$ and it has been conjectured that there exist no more fields K_l

with $h = 1$. H. L. Montgomery and K. Uchida solved this conjecture by proving that the first factor $h_1 > 1$ for $l > 19$ (*Tôhoku Math. J.*, **23** (1971)). J. M. Masley and Montgomery proved that there are precisely 29 distinct K_m with class number 1 (*J. Reine Angew. Math.*, **286/287** (1976)).

According to Kummer, l divides h if and only if l divides h_2 . Since h_1 can be computed explicitly, we can readily determine whether $l \mid h$ or not (- 145 Fermat's Problem; Appendix B, Table 4.111). A prime l is called **regular** if $l \nmid h$, otherwise it is called **irregular**. Thus an odd prime $l \geq 5$ is irregular if and only if there exists an even integer j with $2 \leq j \leq l-3$ such that l divides the numerator of the j th Bernoulli number B_j . This criterion of Kummer can be strengthened as follows. Let S denote the l -Sylow subgroup of the ideal class group \mathfrak{C}_{K_l} of the prime cyclotomic field $K_l = Q(\zeta_l)$. Because the Galois group G of $Q(\zeta_l)/Q$ operates on S , the Abelian group S becomes naturally a $\mathbf{Z}_l[G]$ -module. Choosing the canonical character $\theta: G \rightarrow \mathbf{Z}_l^*$ defined by $\zeta_l^\sigma = \zeta_l^{\theta(\sigma)}$ for any $\sigma \in G$, we have a direct decomposition $S = \prod_{i=0}^{l-2} S^{(i)}$, where $S^{(i)} = \{s \in S \mid s^\sigma = s^{\theta^i(\sigma)}, \sigma \in G\}$. Then it is necessary and sufficient for $l \mid h$ (the numerator of B_j) that $S^{(l-j)} \neq 0$ (K. A. Ribet, *Inventiones Math.*, **34** (1976)). Moreover, because G is naturally isomorphic to the multiplicative group of the reduced residue classes of $\mathbf{Z}/l\mathbf{Z}$, each θ^i becomes a \dagger Dirichlet character. Let

$$B_{1, \theta^{-i}} = \frac{1}{l} \sum_{a \in \mathbf{F}_l^*} \theta^{-i}(a) a;$$

then this number, regarded as an element in Q_l , is equal to a factor of the product (up to ± 1 and 2) appearing in the class number formula for h . Let m_i be the l -exponent of $B_{1, \theta^{-i}}$ for each odd i ($2 < i \leq l-3$). Then it is conjectured that the order of $S^{(i)}$ is precisely equal to l^{m_i} for each odd i with $2 < i \leq l-3$. In particular, if $S' = \prod_{j: \text{even}} S^{(j)}$ is trivial, then this is known to be true. In general, it is also conjectured that $l \nmid h_2$ holds for any prime l (**Vandiver's conjecture**). If the group $S^{(i)}$ is cyclic, then $S^{(i)}$ is of order l^{m_i} (A. Wiles, *Inventiones Math.*, **58** (1980)).

When a Galois extension K over a finite algebraic number field k has the Galois group isomorphic to the additive group of the l -adic integer ring \mathbf{Z}_l , the extension K/k is called a **\mathbf{Z}_l -extension** or **r -extension**. Then for each integer $n \geq 0$ there exists a unique subfield k_n of K with degree l^n over k . Let $\mathbf{k} = K_l = Q(\zeta_l)$ ($l > 2$), and let K be the union of all l^{n+1} th cyclotomic fields $K_{l^{n+1}}$ ($n \geq 0$). Then K/k is a typical example of a \mathbf{Z}_l -extension with subfields $k_n = K_{l^{n+1}}$. Let K/k be a \mathbf{Z}_l -extension with sub-

fields k_n , and let l^{e_n} denote the l -component of the class number of k_n . Then there exist integers $\lambda \geq 0$, $\mu \geq 0$, v independently of n such that $e_n = \lambda n + \mu l^n + v$ for all sufficiently large numbers n (K. Iwasawa, Bull. Amer. Math. Soc., 65 (1959)). These numbers λ , μ , v are called the **Iwasawa invariants** for the Z_l -extension K/k . There exists a unique Z_l -extension K° over the rationals \mathbf{Q} . An extension $K^\circ k$ over k obtained by taking the composite of K° and k is called the **basic Z_l -extension** of k or the **cyclotomic Z_l -extension** of k . We denote its Iwasawa invariants by $\lambda_l(k)$, $\mu_l(k)$, $v_l(k)$. It is known that there are Z_l -extensions with arbitrarily large μ . But it is conjectured that $\mu_l(k) = 0$ holds for the basic Z_l -extension of any k . In particular, in the case $k = K_l = \mathbf{Q}(\zeta_l)$ it has been computed that $\mu_l(k) = 0$ for $l < 125,000$ (for $l < 4001$ by K. Iwasawa and C. C. Sims, J. Math. Soc. Japan, 18 (1966); for $l < 30,000$ by W. Johnson, Math. Comp., 29 (1975); for $l < 125,000$ by S. Wagstaff, Math. Comp., 32 (1978)). By applying the theory of uniform distributions to l -adic situations, B. Ferrero and L. C. Washington (Ann. Math., 109 (1978)) proved that $\mu_l(k) = 0$ when k/\mathbf{Q} is Abelian. In particular, when $k = K$, and $k_n = K_l^{n+1}$, $e_n > 0$ if and only if the class number of K_l is divisible by l (P. Furtwangler, 1911).

Since any quadratic field is a subfield of a cyclotomic field (by a † Gaussian sum formula we have $\mathbf{Q}(\sqrt{m}) \subset \mathbf{Q}(\zeta_{|d|})$, where d is the discriminant of $\mathbf{Q}(\sqrt{m})$), the computation of the class number of quadratic fields and the proof of the law of reciprocity for the † Legendre symbol follow from the arithmetic of cyclotomic fields.

M. Arithmetic of Kummer Extensions

Assume that an algebraic number field k contains an n th primitive root of unity. Then a † Kummer extension $K = k(\sqrt[n]{\mu})$ ($\mu \in k$) is a † cyclic extension of k . Assume that $[K:k] = n$. In order that a prime ideal \mathfrak{p} of k ramify for K/k , it is necessary that \mathfrak{p} be a divisor appearing in (n) or (μ) . If $\mathfrak{p} \nmid (n)$ and $v_{\mathfrak{p}}(\mu) \not\equiv 0 \pmod{n}$, then \mathfrak{p} ramifies for K/k . A prime ideal \mathfrak{p} which is relatively prime to (μ) has the decomposition $E(\mathfrak{p}) = \mathfrak{P}_1 \mathfrak{P}_n$ with distinct primes \mathfrak{P}_i in K if and only if the equation $\mu \equiv \xi^n \pmod{\mathfrak{p}^n}$ is satisfied by some $\xi \in \mathfrak{o}$ for any positive integer m . In particular, if $\mathfrak{p} \nmid (n)$ and $v_{\mathfrak{p}}(\mu) = 0$, we have $E(\mathfrak{p}) = \mathfrak{P}_1 \mathfrak{P}_n$ if and only if $\mu \equiv \xi^n \pmod{\mathfrak{p}}$ is solvable in \mathfrak{o} .

If for an element μ of \mathfrak{o}

$$\mu \equiv \xi^n \pmod{\mathfrak{p}}$$

is solvable by some $\xi \in \mathfrak{o}$, μ is said to be a

residue of the n th power modulo \mathfrak{p} . Assume that $\mathfrak{p} \nmid (n)$ and $v_{\mathfrak{p}}(\mu) = 0$. Let f be the minimal positive integer such that μ^f is a residue of the n th power modulo \mathfrak{p} . Then \mathfrak{p} is decomposed in K as $E(\mathfrak{p}) = \mathfrak{P}_1 \mathfrak{P}_g$, and $N_{K/k}(\mathfrak{P}_i) = \mathfrak{p}^f$ ($i = 1, \dots, g$).

N. Power-Residue Symbol

Let $\zeta_n = \exp(2\pi i/n) \in k$, and let \mathfrak{p} be relatively prime to (n) and (α) ($\alpha \in k$). Then for some r we have

$$\alpha^{(N(\mathfrak{p})-1)/n} \equiv \zeta_n^r \pmod{\mathfrak{p}},$$

and we write

$$\zeta_n^r = \left(\frac{\alpha}{\mathfrak{p}} \right)_n$$

This symbol is called the n th **power-residue symbol** (Kummer). Generalizing this definition, we can define the n th power-residue symbol $(\alpha/\mathfrak{b})_n$ for an ideal \mathfrak{b} of k which is relatively prime to the relative discriminant of $k(\sqrt[n]{\alpha})/k$ by using the † Artin symbol $((K/k)/\mathfrak{b})$:

$$(\sqrt[n]{\alpha})^\sigma = \left(\frac{\alpha}{\mathfrak{b}} \right)_n \sqrt[n]{\alpha}, \quad \sigma = \left(\frac{K/k}{\mathfrak{b}} \right).$$

This symbol satisfies

$$\left(\frac{\alpha}{\mathfrak{b}_1 \mathfrak{b}_2} \right)_n = \left(\frac{\alpha}{\mathfrak{b}_1} \right)_n \left(\frac{\alpha}{\mathfrak{b}_2} \right)_n,$$

$$\left(\frac{\alpha_1 \alpha_2}{\mathfrak{b}} \right)_n = \left(\frac{\alpha_1}{\mathfrak{b}} \right)_n \left(\frac{\alpha_2}{\mathfrak{b}} \right)_n$$

if all the symbols are well defined. In particular, α is a residue of the n th power modulo \mathfrak{p} if and only if $(\alpha/\mathfrak{p})_n = 1$. This symbol coincides with the † quadratic residue symbol for $n = 2$, $k = \mathbf{Q}$, and $\mathfrak{p} \neq 2$.

0. Law of Reciprocity for the Power-Residue Symbol

Several formulas concerning the power-residue symbol are known which are similar to that for the quadratic residue symbol (F. G. M. Eisenstein, Kummer, Furtwangler, Takagi, Artin, Hasse). These can be proved by means of Artin's † general law of reciprocity in class field theory (\rightarrow 59 Class Field Theory).

There are many formulas concerning the reciprocity of the power-residue symbol. One of them is as follows: Let $n = l$ be a prime number. Let $\alpha, \beta \in k$ and assume that (i) α is totally positive; (ii) $v_{\mathfrak{p}}(\alpha) \equiv 0 \pmod{l}$ if $v_{\mathfrak{p}}(\beta) \not\equiv 0 \pmod{l}$, and $v_{\mathfrak{p}}(\beta) \equiv 0 \pmod{l}$ if $v_{\mathfrak{p}}(\alpha) \not\equiv 0 \pmod{l}$, for any prime ideal \mathfrak{p} ; and (iii) $\alpha \equiv 1 \pmod{l}$ and $\beta \equiv 1 \pmod{(1 - \zeta_l)}$. Then

$$\left(\frac{\alpha}{\beta}\right)_l \left(\frac{\beta}{\alpha}\right)_l^{-1} = \zeta_l^a,$$

$$a = \text{Tr}_{k/\mathbb{Q}} \left(\frac{\alpha-1}{l} \frac{\beta-1}{1-\zeta_l} \right)$$

(general law of reciprocity, Hasse, 1924). This result is a generalization of the formula of Eisenstein (1850). If $\alpha(\in k)$ is totally positive and $\alpha \equiv 1 \pmod{l}$, then

$$\left(\frac{\zeta_l}{\alpha}\right)_l = \zeta_l^b, \quad b = \text{Tr}_{k/\mathbb{Q}} \left(\frac{\alpha-1}{l} \right).$$

If $\alpha \equiv 1 \pmod{l(1-\zeta_l)}$, then

$$\left(\frac{l}{\alpha}\right)_l = \zeta_l^c, \quad c = \text{Tr}_{k/\mathbb{Q}} \left(\frac{\alpha-1}{l(1-\zeta_l)} \right);$$

$$\left(\frac{1-\zeta_l}{\alpha}\right)_l = \zeta_l^d, \quad d = -\text{Tr}_{k/\mathbb{Q}} \left(\frac{\alpha-1}{l(1-\zeta_l)} \right)$$

(complementary law of reciprocity, Hasse, 1924).

P. Norm Residue

Let \mathfrak{m} be an integral divisor of k such that $\mathfrak{m} = \prod_i \mathfrak{p}_i^{e_i} \prod_j \mathfrak{p}_j^{(j)} (e_i > 0)$ with finite prime divisors $\{\mathfrak{p}_i\}$ and infinite prime divisors $\{\mathfrak{p}_j^{(j)}\}$, and let β be an element of k that is relatively prime to \mathfrak{m} . For a relative algebraic number field K/k and an element B of K , we set $\beta \equiv N_{K/k}(B) \pmod{\mathfrak{m}}$ if the following two conditions are satisfied: (i) $\beta \equiv N_{K/k}(B) \pmod{\mathfrak{p}_i^{e_i}}$ for every finite \mathfrak{p}_i and (ii) $\beta^{(j)} > 0$ for every infinite prime $\mathfrak{p}_j^{(j)}$ such that $\beta^{(j)}$ is real and its extension to K is imaginary. $\beta \in k$ is then said to be a **norm residue** modulo \mathfrak{m} for K/k if there exists a number B of K such that $\beta \equiv N_{K/k}(B) \pmod{\mathfrak{m}}$.

Let \mathfrak{p} be a finite prime divisor of k . If β is a norm residue modulo \mathfrak{p}^c for a sufficiently large c , then β is a norm residue modulo \mathfrak{p}^e for any $e > c$. Let c be the smallest such integer ($c \geq 0$). Then the ideal $\mathfrak{f}_{\mathfrak{p}} = \mathfrak{p}^c$ is said to be the **\mathfrak{p} -conductor** of norm residue for K/k . If \mathfrak{p} is unramified for K/k , then $c = 0$; i.e., every $\beta \in k$ which is relatively prime to \mathfrak{p} is a norm residue modulo \mathfrak{p}^e for any $e > 0$. For a ramified \mathfrak{p} put $\mathfrak{f}_{\mathfrak{p}} = \mathfrak{p}^c$. For a Galois extension K/k , c is not greater than

$$\sum_{\rho=0}^{r-1} \frac{|V_{\rho}|}{|V_0|} (v_{\rho+1} - v_{\rho}) = \sum_{i=0}^{v_r} \frac{|V^{(i)}|}{|V_0|}.$$

In particular, for an Abelian extension K/k this value is an integer and is equal to c (Hasse, *J. Fac. Sci. Univ. Tokyo*, 1934). For example, the l -conductor of the cyclotomic field $K_{\mathfrak{p}^h}/\mathbb{Q}$ is l^h . We define the \mathfrak{p}_{∞} -conductor for K/k for an infinite prime divisor \mathfrak{p}_{∞} of k by $\mathfrak{f}_{\mathfrak{p}_{\infty}} = \mathfrak{p}_{\infty}$ if \mathfrak{p}_{∞} is real and its extension to K is imaginary, and $\mathfrak{f}_{\mathfrak{p}_{\infty}} = 1$ otherwise (- 257 Local Fields F).

Q. Norm Residue Symbol

For an Abelian extension K/k , the positive divisor

$$\mathfrak{f} = \prod_{\mathfrak{p}} \mathfrak{f}_{\mathfrak{p}}$$

(where \mathfrak{p} runs over all finite and infinite prime divisors of k) is called the **conductor** of K/k (\rightarrow 59 Class Field Theory). For $a \in k$ ($a \neq 0$) take α_0 such that $\alpha/\alpha_0 \equiv 1 \pmod{\mathfrak{f}_{\mathfrak{p}}}$ and $\alpha_0 \equiv 1 \pmod{\mathfrak{f}\mathfrak{f}_{\mathfrak{p}}^{-1}}$, and put $(\alpha_0) = \mathfrak{p}^a \mathfrak{b}$ with \mathfrak{b} relatively prime to \mathfrak{p} . Then \mathfrak{b} is relatively prime to the relative discriminant $\mathfrak{d}_{K/k}$. We define a new symbol by

$$\left(\frac{\alpha, K/k}{\mathfrak{p}}\right) = \left(\frac{K/k}{\mathfrak{b}}\right) \quad (\in G),$$

where $((K/k)/\mathfrak{b})$ is the Artin symbol. This value is independent of the choice of the auxiliary element α_0 . The new symbol is called the **norm-residue symbol** (Hasse, *J. Reine Angew. Math.*, 162 (1930)). In particular, for an infinite real prime divisor $\mathfrak{p}_{\infty}^{(j)}$ of k whose extension $\mathfrak{P}_{\infty}^{(j)}$ for K is imaginary, we have

$$\left(\frac{\alpha, K/k}{\mathfrak{p}_{\infty}^{(j)}}\right) = 1 \text{ or } \sigma$$

according to whether the conjugate $\alpha^{(j)}$ is positive or negative, where σ is the automorphism of K/k induced from the complex conjugation of the completion C of K with respect to $\mathfrak{P}_{\infty}^{(j)}$.

The norm-residue symbol has the following properties:

$$(1) \left(\frac{\alpha\alpha', K/k}{\mathfrak{p}}\right) = \left(\frac{\alpha, K/k}{\mathfrak{p}}\right) \left(\frac{\alpha', K/k}{\mathfrak{p}}\right);$$

(2) if \mathfrak{p} is unramified for K/k , then

$$\left(\frac{\alpha, K/k}{\mathfrak{p}}\right) = \left(\frac{K/k}{\mathfrak{p}}\right)^{-v_{\mathfrak{p}}(\alpha)};$$

(3) in order that α be a norm residue modulo $\mathfrak{f}_{\mathfrak{p}}$ for K/k , it is necessary and sufficient that

$$\left(\frac{\alpha, K/k}{\mathfrak{p}}\right) = 1;$$

(4) the **product formula** for the norm-residue symbol (Hasse) is

$$\prod_{\mathfrak{p}} \left(\frac{\alpha, K/k}{\mathfrak{p}}\right) = 1,$$

where \mathfrak{p} runs over all finite and infinite prime divisors of k ; and (5) if the domain of the variable α is the whole k ($\neq 0$), or the set of all α such that (α) is relatively prime to \mathfrak{p} , or the set of all α such that $\alpha \equiv 1 \pmod{\mathfrak{p}^n}$ ($v_{\mathfrak{p}} + 1 \leq m \leq v_{\mathfrak{p}+1}$), then the range of $((\alpha, K/k)/\mathfrak{p})$ is the decomposition group Z of \mathfrak{p} , the inertia group T of \mathfrak{p} , or the ramification group V_{ρ}

of p , respectively (Hasse, S. Iyanaga, 1933) (- 257 Local Fields F).

R. Hilbert Norm-Residue Symbol

The symbol first introduced by Hilbert for quadratic fields can be defined in a general algebraic number field k containing an n th primitive root ζ_n of unity. Let $\alpha, \beta \in k$ ($\alpha \neq 0$, $\beta \neq 0$), and let p be a prime divisor. Then the following n th root of unity $((\alpha, \beta)/p)_n$ is defined by using the norm-residue symbol:

$$\left(\frac{\alpha, \beta}{p}\right)_n \sqrt[n]{\beta} = (\sqrt[n]{\beta})^\sigma,$$

$$\sigma = \left(\frac{\alpha, k(\sqrt[n]{\beta})/k}{p}\right).$$

This symbol $((\alpha, \beta)/p)_n$ is called the **Hilbert norm-residue symbol**. It is also called the **Hilbert-Hasse norm-residue symbol**. For $\alpha, \alpha', \beta \in k$, we have

$$(1) \left(\frac{\alpha\alpha', \beta}{p}\right)_n = \left(\frac{\alpha, \beta}{p}\right)_n \left(\frac{\alpha', \beta}{p}\right)_n;$$

(2) the law of symmetry:

$$\left(\frac{\alpha, \beta}{p}\right)_n = \left(\frac{\beta, \alpha}{p}\right)_n^{-1};$$

and (3) the **product formula** for the Hilbert norm-residue symbol:

$$\prod_p \left(\frac{\alpha, \beta}{p}\right)_n = 1$$

(Hilbert, Furtwängler, Takagi, Artin, Hasse). For detailed properties concerning the norm-residue symbol, power-residue symbol, and Hilbert norm-residue symbol and for references for them see Hasse [6].

In general, the problem of obtaining various laws of reciprocity for the power-residue symbol is reduced to the one of computing the Hilbert norm-residue symbol explicitly. Detailed formulas for these symbols are called **explicit reciprocity laws**; they are treated as a topic in the number theory of local fields (- 257 Local Fields H).

S. Density Theorem

Let M be a set of prime ideals of k . If

$$\lim_{s \rightarrow 1+0} \sum_{p \in M} \frac{1}{N(p)} \bigg/ \log \frac{1}{s-1}$$

exists, its value is said to be the **density** of M . The density of the set of all prime ideals of k is 1. Let H be an ideal group modulo an integral divisor m . Then the density of the set of all prime ideals contained in each coset of $\mathfrak{J}(m)$

modulo H is $1/|\mathfrak{J}(m):H|$. In particular, let H be the ray $S(m)$. Then this result implies that each coset of $\mathfrak{J}(m)$ modulo $S(m)$ contains infinitely many prime ideals (a generalization to algebraic number fields of the **†prime number theorem** for arithmetic progression).

Let K/k be a Galois extension, C be a conjugate class of the Galois group G of K/k , and $M(C)$ be the set of all prime ideals p of k such that the **†Frobenius** automorphism of each prime factor \mathfrak{P}_i of p in K belongs to C . Then the density of $M(C)$ is $|C|/|G|$ (**Chebotarev's density theorem**, *Math. Ann.*, 95 (1926)).

Each element σ of the Galois group G of K/k can be expressed by the permutation z of the conjugate fields $K^{(1)}, \dots, K^{(n)}$ of K over k . Let z be expressed as the product of r **†cycles** of length f_1, \dots, f_r . Hence $n = f_1 + \dots + f_r$. Let $C(f_1, \dots, f_r)$ be the set of all such σ in G , and let $M(f_1, \dots, f_r)$ be the set of all prime ideals p of k such that p is decomposed in K/k as the product of r prime ideals of K with relative degree f_1, \dots, f_r . Then the density of $M(f_1, \dots, f_r)$ is $|C(f_1, \dots, f_r)|/|G|$ (Artin, *Math. Ann.*, 89 (1923)).

T. Relation to the Arithmetic of Local Fields

It is quite useful to investigate the relation between the arithmetic of algebraic number fields and that of local fields. For example, let a prime ideal p of an algebraic number field k be decomposed as $E(p) = \mathfrak{P}_1^{e_1} \cdot \mathfrak{P}_2^{e_2} \cdot \dots \cdot \mathfrak{P}_g^{e_g}$, $N_{K/k}(\mathfrak{P}_i) = p^{f_i}$ ($i = 1, \dots, g$) in an extension K of k . Let $K_{\mathfrak{p}}$ and k_p be the completion of K and k with respect to \mathfrak{P} -adic and p -adic valuations, respectively. Then we have $[K_{\mathfrak{p}} : k_p] = e_i f_i$ and $K \otimes_k k_p \cong K_{\mathfrak{p}_1} + \dots + K_{\mathfrak{p}_g}$ (direct sum). The relative different $\mathfrak{D}_{K/k}$ is expressed as (the p -component of $\mathfrak{D}_{K/k}$) $= \prod_{i=1}^g \mathfrak{D}_{K_{\mathfrak{p}_i}/k_p}$. For a Galois extension K/k the p -conductor $\mathfrak{f}_p = p^c$ for the norm-residue and the conductor p^c of local extension $K_{\mathfrak{p}}/k_p$ have the same exponent c . For a local field $K_{\mathfrak{p}}/k_p$, each norm-residue modulo \mathfrak{f}_p is a norm of an element of K . Hence precise results concerning the norm-residue in local fields can be applied immediately to a global field K/k (- 257 Local Fields).

We can also apply the method of the idele group of an algebraic number field k , and therefrom we can prove results concerning the ideal class group, unit group, and zeta function of k (\rightarrow 6 Adeles and Ideles).

U. History of the Arithmetic of Algebraic Number Fields

C. F. Gauss (1832) was the first to generalize the notion of integers to algebraic number

fields in considering the elements of $\mathbf{Z}[\sqrt{-1}]$, now called **Gaussian integers** ($\mathbf{Z}[\sqrt{-1}]$ is the principal order of $\mathbf{Q}(\sqrt{-1})$). After investigations by G. L. Dirichlet and Kummer, the notion of ideals was introduced by Dedekind (1871) [2]. L. Kronecker gave another foundation for the arithmetic of algebraic number fields (1882) [3]. Dirichlet proved the unit theorem and, introducing the analytic method to number theory, gave the class number formula of quadratic fields (\rightarrow 347 Quadratic Fields). H. Minkowski first applied the theory of lattice points to number theory (\rightarrow 182 Geometry of Numbers), and K. Hensel introduced the p -adic method (\rightarrow 257 Local Fields). Hilbert (1897) [4] and Hasse (1926, 1927, 1930) [6] summarized the main results on the arithmetic of algebraic number fields known at that time. In particular, Hilbert's report centered around the arithmetic of Galois extensions, and Hasse's around the class field theory obtained by T. Takagi, E. Artin, and H. Hasse (\rightarrow 59 Class Field Theory). Since c. 1950, when the notions of ideles and adeles were introduced, cohomology-theoretic methods have been successfully applied to number theory (\rightarrow 6 Adeles and Ideles). Recently various local methods, for example, the Iwasawa theory of \mathbf{Z}_p -extensions, t formal groups, and $^t p$ -adic L -functions (T. Kubota and H. W. Leopoldt, *J. Reine Angew. Math.*, 214/215 (1964)) have been frequently applied to research in algebraic number fields.

References

- [1] K. F. Gauss, *Theoria residuorum biquadraticorum* I (1825), II (1831), Werke 2, Göttingen, 1863, 65–148.
- [2] R. Dedekind, *Dirichlet's Vorlesungen über Zahlentheorie*, fourth edition, 1894, Supplement XI (Gesammelte mathematische Werke 3, Braunschweig, 1932) (Chelsea, 1969).
- [3] L. Kronecker, *Grundzüge einer arithmetischen Theorie der algebraischen Grosse*, *J. Reine Angew. Math.*, 92 (1882), 1–122 (Werke 2, Teubner, 1897, 237–387) (Chelsea, 1968).
- [4] D. Hilbert, *Die Theorie der algebraischen Zahlkörper*, *Jber. Deutsch. Math. Verein.* 4, (1897), 175–546 (Gesammelte Abhandlungen 1, Springer, 1932, 63–363) (Chelsea, 1967).
- [5] E. Hecke, *Vorlesungen über die Theorie der algebraischen Zahlen*, Akademische Verlag, 1923 (Chelsea, 1970).
- [6] H. Hasse, *Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper*, *Jber. Deutsch. Math. Verein.*, 1, 35 (1926), 1–55; Ia, 36 (1927), 233–311; II (1930) (Physica Verlag, 1965).

- [7] H. Weyl, *Algebraic theory of numbers*, Ann. Math. Studies, Princeton Univ. Press, 1940.
- [8] H. Hasse, *Zahlentheorie*, Akademie-Verlag, third edition, 1969; English translation, Number theory, Springer, 1980.
- [9] E. Weiss, *Algebraic number theory*, McGraw-Hill, 1963.
- [10] S. Lang, *Algebraic number theory*, Addison-Wesley, 1970.
- [11] E. Artin, *Algebraic numbers and algebraic functions*, Lecture notes, Princeton Univ., 1950–1951 (Gordon & Breach, 1967).
- [12] S. Iyanaga (ed.), *Theory of numbers*, North-Holland, 1975. (Original in Japanese, 1969.)
- [13] A. Weil, *Basic number theory*, Springer, second edition, 1973.
- [14] J. W. S. Cassels and A. Fröhlich (eds.), *Algebraic number theory*, Academic Press, 1967.
- [15] H. Hasse, *Über die Klassenzahl abelscher Zahlkörper*, Akademie-Verlag, 1952.
- [16] Z. I. Borevich and I. R. Shafarevich, *Number theory*, Academic Press, 1966. (Original in Russian, 1964.)
- [17] S. Lang, *Cyclotomic fields*, Springer, 1978, II, 1980.

15 (VIII.3) Algebraic Surfaces

A. Definition

An algebraic variety of dimension 2 is called an **algebraic surface**. In this article, by a surface we mean a complete irreducible algebraic surface defined over an algebraically closed field \mathbf{K} .

B. History

The history of algebraic surfaces originated with the study of algebraic functions of two variables. In the case of algebraic functions of one variable, the introduction of t Riemann surfaces attached to such functions played an essential role in the development of the theory. The study of algebraic functions of two variables led naturally to the consideration of the surfaces defined by a suitable polynomial equation. H. Poincaré and E. Picard are among those who studied the homological structure of the surface defined by the equation $P(x, y, z) = 0$. The theory of t Abelian integrals (Picard integrals) is one of the consequences of

such topological investigations. S. Lefschetz obtained further results in this direction.

M. Noether and geometers of the Italian school, such as F. Enriques, G. Castelnuovo, and F. Severi, studied algebrogeometric properties of algebraic surfaces. In particular, the Italian school geometers recognized the importance of irregularity and thoroughly investigated its geometric meaning. In the early 20th Century they succeeded in constructing the great edifice of the theory of algebraic surfaces. Though some of their results lack rigorous proof, efforts to build a foundation for those results have led to the recent developments in algebraic geometry. A significant contribution to the modernization of the theory was made by O. Zariski and K. Kodaira.

The resolution of singularities of an algebraic surface is one of the most fundamental problems in the field. When the universal domain is the complex number field, function-theoretic methods were used by Italian-school geometers and R. J. Walker (Ann. Math., 36 (1935)). Zariski introduced the valuation-theoretic method to deal with the problem when the characteristic of the universal domain is zero. S. Abhyankar (1966) succeeded in resolving the case of positive characteristics.

C. Divisors and Linear Systems

In what follows, let S denote a nonsingular surface. S can be embedded into some projective space. Let Σ be a linear system of divisors on S and f_0, \dots, f_n be a basis of the defining module for Σ over \mathbf{K} . Associated with Σ we have a rational mapping $\Phi_\Sigma: S \rightarrow \mathbf{P}^n$ defined by $\Phi_\Sigma(P) = (f_0(P), \dots, f_n(P))$ for general points P on S . Pullbacks of hyperplanes by Φ_Σ are called **variable components** of Σ . Any member of Σ is a sum of a variable component and a fixed component of Σ . Let S' denote the closure of the image of S by Φ_Σ . If $\dim S' = 2$, a general variable component is irreducible. If $\dim S' = 1$, then a general variable component is composed of an algebraic system of dimension 1, which is called an **algebraic pencil**. These result from Bertini's theorems. For any divisors D and D' on S , the **intersection number** (or the Kronecker index) $I(D \cdot D')$ is defined; this number is a symmetric bilinear form such that $I(D \cdot D) = I(D_1 \cdot D'_1)$ for any divisors D_1 and D'_1 linearly equivalent to D and D' , respectively. If C is a nonsingular curve on S , $I(C \cdot D)$ coincides with the degree of the restriction $C \cdot D$ of D to C . $I(D \cdot D) = (D^2)$ is said to be the **self-intersection number** of D . If D is an ample divisor on S , then $(D^2) > 0$ and $I(D \cdot C) > 0$ for any irreducible curve C on S . These properties

characterize an ample divisor (Y. Nakai; \rightarrow 16 Algebraic Varieties E).

Let Σ be an irreducible linear system of dimension $r (\geq 1)$, and let C be a generic component of Σ . Let C' be a member of Σ different from C . Then the set of C -divisors $C \cdot C'$ forms a linear system of dimension $r - 1$ on C . This is called the **trace** of Σ on C and is denoted by $\text{Tr}_C \Sigma$. The trace is, in general, not complete. The integer $\dim |\text{Tr}_C \Sigma| - \dim \text{Tr}_C \Sigma = \delta(\Sigma)$ is called the **deficiency** of Σ . The deficiency of the complete linear system $|D|$ is denoted by $S(D)$.

Let x be a point of S , and let \mathcal{O}_x be the local ring of x . Then $\mathcal{O}_S = \bigcup_{x \in S} \mathcal{O}_x$ is an algebraic coherent sheaf, called the **structure sheaf** of S . Let D be a divisor on S . The sheaf of germs of rational functions f such that $(f) + D > 0$ is denoted by $\mathcal{G}'(D)$. Then $H^0(S, \mathcal{O}_S(D))$ is a defining module for the complete linear system $|D|$. $D > 0$ if and only if $\mathcal{O}_S(-D)$ is a sheaf of \mathcal{O}_S -ideals. The quotient sheaf $\mathcal{O}_S/\mathcal{O}_S(-D)$ will be denoted by \mathcal{O}_D . If D is a prime divisor, then \mathcal{O}_D is the structure sheaf of the algebraic curve D (\rightarrow 9 Algebraic Curves). Let \mathcal{F} be a sheaf on S . We set

$$\chi(S, \mathcal{F}) = \sum_{q=0}^2 (-1)^q \dim H^q(S, \mathcal{F}).$$

$\chi(S, \mathcal{O}_S)$ will be denoted simply by $\chi(S)$. We call $p_a(S) = \chi(S) - 1$ the **arithmetic genus** of the algebraic surface S . Sometimes $\chi(S)$ is referred to as the arithmetic genus of S . We set $\chi_S(D) = \chi(S) - \chi(S, \mathcal{O}_S(-D))$. The integer $p_r(D) = 1 - \chi_S(D)$ is, by definition, the **arithmetic genus of the divisor** D . If D is a prime divisor, then $p_r(D)$ coincides with the arithmetic genus of the algebraic curve, i.e., $p_r(D) = \dim H^1(D, \mathcal{O}_D)$.

D. Riemann-Roch Theorem

Let S be a nonsingular surface and let K denote a canonical divisor on S , i.e., $K = (\omega)$ for some nonzero rational 2-form ω on S . If C is a nonsingular irreducible curve on S , $(K + C) \cdot C$ becomes a canonical divisor on C ; hence, $\deg((K + C) \cdot C) = 2p_a(C) - 2$ by a corollary to the Riemann-Roch theorem on C (\rightarrow 9 Algebraic Curves C). Since $I((K + C) \cdot C) = \deg((K + C) \cdot C)$, it follows that $p_a(C) = I((K + C) \cdot C)/2 + 1$. Moreover, the formula

$$p_a(D) = I((K + D) \cdot D)/2 + 1$$

holds for an arbitrary curve D on S ; this is called the **adjunction formula**. For any divisor, we define $p_r(D)$ to be $I(D \cdot (D + K))/2 + 1$. Then

$$\begin{aligned} \chi(S, \mathcal{O}(D)) &= \chi(S) - \chi_S(-D) = \chi(S) + p_a(-D) - 1 \\ &= I(D \cdot (D - K))/2 + \chi(S). \end{aligned}$$

This formula is called the **Riemann-Roch theorem** on S (\rightarrow 366 Riemann-Roch Theorems C). Applying Serre's duality theorem to D (\rightarrow 16 Algebraic Varieties E) we have

$$\dim H^i(S, \mathcal{O}(D)) = \dim H^{2-i}(S, \mathcal{O}(K-D))$$

for $i=0, 1, 2$.

In particular, $\dim H^2(S, \mathcal{O}(D)) = l(K-D)$, which is called the **index of speciality** of D ; $\dim H^1(S, \mathcal{O}(D)) = \dim H^1(S, \mathcal{O}(K-D))$ is called the **superabundance** of D , denoted by $h^1(D)$. The inequality

$$l(D) + l(K-D) \geq \chi(S, \mathcal{O}(D)) = I(D \cdot (D-K))/2 + \chi(S)$$

is called the **Riemann-Roch inequality**, where equality holds if and only if $h^1(D) = 0$. If D is a curve with s connected components, then

$$l(D+K) = h^1(-D) + p_a(D) - s + \chi(S).$$

In addition, if $h^1(0) = 0$, then $h^1(-D) = 0$ and thus $l(D+K) = p_a(D) - s + \chi(S)$. This is called the **Riemann-Roch theorem for the adjoint system**. Note that $|D+K|$ is called the **adjoint system** of D . The Noether formula, $(K^2) + c_2(S) = 12\chi(S)$, is a special case of Hirzebruch's theorem of Riemann-Roch type (\rightarrow 366 Riemann-Roch Theorems B). Here, $c_2(S)$ denotes the second Chern number of S , which coincides with the Euler number of S if \mathbf{K} is the field of complex numbers.

Let $\text{Div}(S)$ denote the group of all divisors on S ; by linearity we can define the bilinear form $I(D \cdot D')$ on $\text{Div}(S)_0 = \text{Div}(S) \otimes_{\mathbf{Z}} \mathbf{Q}$. $J = \{D \in \text{Div}(S)_0 \mid I(D \cdot D') = 0 \text{ for all } D'\}$ is a subgroup and $X = \text{Div}(S)_0/J$ is a finite-dimensional vector space over \mathbf{Q} , on which the nondegenerate bilinear form I is induced. I has a unique positive eigenvalue; thus the other eigenvalues are all negative. This is called the **index theorem of Hodge**; it is derived from the Riemann-Roch theorem on S . From this, we infer that if $D^2 = I(D \cdot D) > 0$, then $I(D \cdot D')^2 \geq D^2 \cdot D'^2$ for any D' . X is said to be the **Neron-Severi group** of S and $\dim X$ is said to be the **Picard number** of S (\rightarrow 16 Algebraic Varieties P).

To study $l(mD)$ as a function of $m \gg 0$, O. Zariski writes an arbitrary effective divisor D as a sum of $D^{(+)}$ and $D^{(-)} \in \text{Div}(S)_0$ with non-negative rational coefficients such that (1) $D^{(+)}$ is **arithmetically effective** (or, numerically semipositive), i.e., $I(D^{(+)} \cdot C) \geq 0$ for any curve C on S ; (2) $D^{(-)} = 0$ or the intersection matrix of the support of $D^{(-)}$ is negative definite; and (3) $I(D^{(+)} \cdot D^{(-)}) = 0$. Such a decomposition is unique and is called the **Zariski decomposition** of D [7]. If $mD^{(+)}$ is a divisor for some $m > 0$, then $l(mD) = l(mD^{(+)})$.

E. Invariants of Algebraic Surfaces

There are many invariants besides the arithmetic genus discussed earlier. We set $h^{q,p} = \dim H^p(S, \Omega^q)$. Then $h^{2,0}$ is equal to the number of linearly independent holomorphic 2-forms; it is called the **geometric genus** and is usually denoted by p_g . Since $h^{0,1}$ gives the maximum among the deficiencies $\delta(\Sigma)$ of linear systems on S , it is called the **maximal deficiency** of S . For a divisor C such that $h^1(C) = 0$, we have $\delta(C) = h^{0,1}$. The number $h^{0,1}$ was formerly called the **irregularity** of S , because $h^{0,1}$ was considered to be a correction term in the equality $p_a(S) = p_g - h^{0,1}$. The study of higher-dimensional varieties showed, however, that it was unnatural to regard $h^{0,1}$ as a correction term. At present, by **irregularity** we mean the dimension of the Picard variety of S (\rightarrow 16 Algebraic Varieties P), and we denote this number by q .

When S is defined over the complex number field, we have $h^{p,q} = h^{q,p}$. In particular, $q = h^{0,1} = h^{1,0}$. This number is equal to the number of linearly independent \dagger Abelian simple integrals of the first kind; it is also equal to one-half the first Betti number of S . In cases with positive characteristic, these equalities do not hold in general. J.-P. Serre gave an example of an algebraic surface S such that $h^{0,1} \neq h^{1,0}$, and J. Igusa gave an example such that $q < h^{0,1} = h^{1,0}$. Let K be a canonical divisor of S . The number $P_i = l(iK)$ is called the **i -genus**, and P_i ($i = 2, 3, \dots$) are generally called **plurigenera**. If $P_n = 0$ and $d \mid n$, then P_d is also zero. The numbers $p_a(S) = h^{2,0} - h^{0,1}$, $p_g(S) = h^{2,0}$, $h^{1,0}$, $h^{0,1}$, P_i ($i = 2, 3, \dots$) ($P_1 = p_g$) are \dagger absolute invariants of S ; i.e., they take the same values for any nonsingular surface S' that is birationally equivalent to S . However, $h^{1,1}$ is not an absolute invariant. For a projective plane, all plurigenera P_n vanish and the irregularity $q = 0$. Thus if S is a **rational surface** (\rightarrow 16 Algebraic Varieties J), i.e., a surface which is birationally equivalent to \mathbf{P}^2 , all $P_n = q = 0$. Conversely, any surface with $q = P_2 = 0$ is a rational surface. This is called **Castelnuovo's criterion**. A **ruled surface** is defined to be a surface that is birationally equivalent to a product of the projective line and a curve. All P_n of a ruled surface equal 0 and any surface with $P_4 = P_6 = 0$ is a ruled surface. This is called the **criterion of ruled surfaces** (Enriques).

F. Characteristic Linear Systems of Algebraic Families

One of the central problems considered by the Italian school was to prove that the irregular-

ity q is equal to the maximal deficiency $h^{0,1}$. For that purpose, Severi introduced the notion of characteristic linear systems of algebraic families. Let Σ be an irreducible algebraic family of positive divisors on S such that a generic member C of Σ is an irreducible nonsingular curve, and let r be the dimension of Σ . Let Σ_1 be a 1-dimensional subfamily of Σ containing C as a simple member, and let C be a generic member of Σ_1 . Then the specialization of $C' \cdot C$ over the specialization $C' \rightarrow C$ is a well-defined C -divisor of degree $\eta = I(C \cdot C')$. The set of C -divisors thus obtained is called the **characteristic set**. The characteristic set forms an $(r-1)$ -dimensional linear system and contains $\text{Tr}_C[C]$ as a subfamily. This linear system is called the **characteristic linear system** of Σ . For any algebraic family of dimension r , we have $r \leq \dim C| + q$. In particular, there exists an algebraic family Σ that contains ∞^q linear systems and such that for a generic curve C we have $h^1(C) = 0$. For such an algebraic family, we have the equality $r = \dim C| + q$; hence the inequality $q \leq h^{0,1}$ follows. Moreover, if the characteristic linear system is complete, we have $q = h^{0,1}$. The proof of the completeness of characteristic linear systems given by Severi is valid only in some special cases (e.g., the case $p_g = 0$). For a complex algebraic manifold, a rigorous proof was given later. When the characteristic is positive, the completeness does not hold in general (Igusa); however, for the surface with $p_g = 0$, the completeness holds (Y. Nakai). The completeness holds if and only if the Picard scheme of S is reduced [14].

G. Birational Transformations of Algebraic Surfaces

Let S and S' denote nonsingular surfaces. If there exists a birational morphism $T: S \rightarrow S'$, we say that S dominates S' , and we write $S \geq S'$. In addition, if T is not an isomorphism, we write $S > S'$. In case there does not exist an S' with $S > S'$, S is said to be **relatively minimal**. On the other hand, if we have $S' > S$ for any S' which is birationally equivalent to S , S is said to be **minimal**. Any minimal surface is, by definition, relatively minimal. If a minimal (resp. relatively minimal) surface S is birationally equivalent to S' , we say that S is a **minimal** (resp. **relatively minimal**) **model** of S' or of the field $\mathbf{K}(S)$. A necessary and sufficient condition for S to have a minimal model is that S not be a ruled surface (Castelnuovo and Enriques).

Let S be a nonsingular surface and P be a point on S . Replacing P by a projective line \mathbf{P}^1 , we have a nonsingular surface S' and a birational morphism $T: S' \rightarrow S$ such that $T^{-1}(P)$

$= E \cong \mathbf{P}^1$ and $S' - E \cong S - \{P\}$ by $T: T: S' \rightarrow S$ (or T^{-1} in some references) is said to be a locally quadratic transformation. Any birational morphism between nonsingular surfaces is a composition of locally quadratic transformations and an isomorphism. Given a birational mapping $T: S \rightarrow S'$ and P on S' , $T^{-1}\{P\}$ is called an **exceptional curve** whenever it is not a point. Moreover, it is called an **exceptional curve of the first kind** if T is regular along $T^{-1}\{P\}$. Otherwise, it is of the **second kind**. Exceptional curves consist of irreducible rational curves. An irreducible curve E is an exceptional curve of the first kind if and only if $(E^2) = -1$ and $E \cong \mathbf{P}^1$. S is relatively minimal if and only if S has no exceptional curves of the first kind; S is minimal if and only if S has no exceptional curves at all. A relatively minimal surface that is not minimal is a ruled surface. Such a surface is either \mathbf{P}^2 or a \mathbf{P}^1 -bundle over a curve. In particular, a relatively minimal rational surface is either \mathbf{P}^2 or a \mathbf{P}^1 -bundle over \mathbf{P}^1 . Any surface of the latter type is occasionally called a **Hirzebruch surface** and its (K^2) is 8, where K denotes a canonical divisor. However, (K^2) of \mathbf{P}^2 is 9. Define the **linear genus** of a rational surface to be 10. If S is not a rational surface, taking a relatively minimal surface S' , we define the **linear genus** $p^{(1)}$ of S to be $(K^2) + 1$ of S' .

H. Examples of Algebraic Surfaces

Let S_m denote a nonsingular surface in \mathbf{P}^3 defined by a homogeneous polynomial of degree m . Let H denote a divisor on S_m induced from a (hyper)plane on \mathbf{P}^3 . Then the canonical divisor K is linearly equivalent to $(m-4)H$, i.e. $K \sim (m-4)H$. Hence, $p_g = (m-1)(m-2)(m-3)/6$. Moreover, $q = 0$ if \mathbf{K} is the field of complex numbers, S_m is simply connected. S_2 is isomorphic to the product of two copies of \mathbf{P}^1 , i.e. $S_2 \cong \mathbf{P}^1 \times \mathbf{P}^1$; hence, it is a rational surface. S_3 is also rational. There exist 27 lines on S_3 . Contracting 6 mutually disjoint lines among these 27 lines, we obtain a projective plane. Conversely, given 6 points on \mathbf{P}^2 in general position, by performing locally quadratic transformations with these points as centers we get a cubic surface S_3 and a birational morphism $T: S_3 \rightarrow \mathbf{P}^2$. The inverse images by T of the 6 given points, the proper transforms of 15 lines connecting every pair of points chosen from the 6, and 6 conics passing every 5 of the 6 points by T^{-1} , exhaust 27 lines on S_3 . If $m = 4$, then $K \sim 0$. In general, a nonsingular surface with $K \sim 0$ and $q = 0$ is said to be a **K3 surface** (- 72 Complex Manifolds J). K3 surfaces have certain properties similar to **Abelian surfaces** that are defined to be 2-

dimensional Abelian varieties (\rightarrow 3 Abelian Varieties). For $\mathbf{K} = \mathbf{C}$, the period mapping defined by integrating regular 2-forms has been studied extensively in connection with moduli theory.

In general, surfaces with the bigenus $P_2 = p_g = 1$ and $q = 0$ are birationally equivalent to $K3$ surfaces; surfaces with $P_4 = p_g = 1$ and $q = 2$ are birationally equivalent to Abelian surfaces. Every Abelian surface has the involution i defined by $i(x) = -x$ and the quotient surface by i is a singular surface with 16 ordinary double points. By performing 16 locally quadratic transformations successively with these singular points as centers we obtain a $K3$ surface. Such a $K3$ surface is called a **Kummer surface**. The original singular surface is also called a Kummer surface. A quartic surface in \mathbf{P}^3 with 16 double points is an example of a Kummer surface.

The theory of birational classification of surfaces was developed by Castelnuovo, Enriques, and others of the Italian school. This theory has been extensively enriched and generalized in various ways. Kodaira's theory of analytic surfaces includes classification of algebraic surfaces (\rightarrow 72 Complex Manifolds I, J), and classification of surfaces in the positive-characteristic case has been recently studied in detail [4, 5]. When the field $\mathbf{K}(S)$ is a subfield of a purely transcendental extension $\mathbf{K}(X, Y)$, S is said to be a unirational surface. If the extension $\mathbf{K}(X, Y)/\mathbf{K}(S)$ is separable, S is a rational surface. However, if it is inseparable, S may become nonrational, a $K3$ surface, an elliptic surface, or a surface of general type [5, 12]. Even for noncomplete surfaces, we have a classification theory [10] similar to the previous ones. The following result is one of the applications: whenever $S \times \mathbf{A}^n \cong \mathbf{A}^3$, S is isomorphic to \mathbf{A}^2 . Here \mathbf{A}^n means the affine n -space (M. Miyanishi, T. Sugie, T. Fujita; [7, 9, 13]).

Let I be the ring of integers of a real quadratic field of discriminant d . The Hilbert modular group $G = SL(2, I)/\{\pm 1\}$ acts on the product \mathcal{H}^2 , \mathcal{H} being the complex upper half-plane. The normal complex space \mathcal{H}^2/G can be compactified by adding a finite number of points and thus a compact normal surface is obtained. Resolving these singularities in the canonical minimal way, we have a nonsingular surface $Y(d)$ over \mathbf{C} , which is called the **Hilbert modular surface** with discriminant d . $Y(d)$ is simply connected; hence, $q(Y(d)) = 0$. If $d = 5, 8, 12, 13, 17, 21, 24, 28, 33, 60$, then $Y(d)$ is a rational surface. If $d = 29, 37, 40, 41, 44, 56, 57, 69, 105$, then $Y(d)$ is birationally equivalent to a $K3$ surface. If $d = 53, 61, 65, 73, 76, 77, 85, 88, 92, 93, 120, 140, 165$, then $Y(d)$ is an elliptic surface with $\kappa(Y(d)) = 1$. Otherwise, it is a

surface of general type (\rightarrow [19, 201 in 72 Complex Manifolds]).

For O -cycles on a surface \rightarrow 16 Algebraic Varieties J.

References

- [1] M. Artin, Some numerical criteria for contractibility of curves on algebraic surfaces, *Amer. J. Math.*, 84 (1962), 4855-496.
- [2] H. F. Baker, Introduction to the theory of algebraic surfaces and higher loci, *Principles of Geometry* 6, Cambridge Univ. Press, 1933.
- [3] A. Beauville, Surfaces algebriques complexes, *C.I.M.E.* (1977), *Astérisque* 54 (1978).
- [4] E. Bombieri and D. Husemoller, Classification and embeddings of surfaces, *Algebraic Geometry*, Arcata, 1974, *Amer. Math. Soc. Proc. Symp. Pure Math.*, 29 (1975), 3299-420.
- [5] E. Bombieri and D. Mumford, Enriques classification of surfaces in characteristic p : II, *Complex Analysis and Algebraic Geometry* (In Honour of Kodaira), Iwanami Shoten and Cambridge Univ. Press (1977), 233-42, III, *Invent. Math.*, 35 (1976), 197-232.
- [6] F. Enriques, *Le superficie algebriche*, Zanichelli, 1949.
- [7] T. Fujita, On Zariski problem, *Proc. Japan Acad.*, 55 (1979), 106-110.
- [8] H. W. E. Jung, *Algebraische Flächen*, Helwingsche, 1925.
- [9] T. Kambayashi, On Fujita's strong cancellation theorem for the affine plane, *J. Fac. Sci. Univ. of Tokyo*, 27 (1980), 535-548.
- [10] Y. Kawamata, On the classification of non-complete algebraic surfaces, *Proc. Algebraic Geometry*, Copenhagen, 1978, *Lecture notes in math.* 732, Springer, 1979, 215-232.
- [11] K. Kodaira, *Collected Works* III, Iwanami, 1975.
- [12] M. Miyanishi, Unirational quasi-elliptic surfaces, *Japan. J. Math.*, 3 (1977), 395-416.
- [13] M. Miyanishi and T. Sugie, Affine surfaces containing cylinderlike open sets, *J. Math. Kyoto*, 20 (1980), 11-42.
- [14] D. Mumford, Lectures on curves on an algebraic surface, *Ann. Math. Studies*, Princeton Univ. Press, 1966.
- [15] D. Mumford, Enriques' classification of surfaces in char p : I, *Global Analysis* (In Honor of Kodaira), Univ. of Tokyo Press and Princeton Univ. Press, 1969.
- [16] M. Nagata, On rational surfaces I, 11, *Mem. Coll. Sci. Univ. Kyôto*, (A) 32 (1960), 351-370; 33 (1961), 271-293.
- [17] E. Pascal, *Repertorium der höheren Mathematik*, II., Teubner, second edition, 1922.
- [18] E. Picard and G. Simart, *Théorie des*

fonctions algébriques de deux variables indépendentes I, II, Gauthier-Villars, 1897, 1906.

[19] A. N. Rudakov and I. R. Safarevich, Inseparable morphisms of algebraic surfaces, Math. USSR-I-V., 10 (1976), 1205-1237.

[20] I. R. Safarevich, Algebraic surfaces, Proc. Steklov Inst. Math., Amer. Math. Soc., 1967.

[21] O. Zariski, Collected Papers II, III, MIT Press, 1973, 1978.

[22] O. Zariski, Algebraic surfaces, Erg. Math., Springer, second edition, 1971.

16 (VIII.4) Algebraic Varieties

A. Affine Algebraic Varieties and Projective Algebraic Varieties

Fix a field k . A subset of the n -dimensional \dagger affine space k^n over k is called an **affine algebraic variety** (or simply **affine variety**) if it can be expressed as the set of the common zeros of a (finite or infinite) set of polynomials $F_i(X_1, \dots, X_n)$ with coefficients in k . Similarly, a subset of the n -dimensional \dagger projective space $P^n(k)$ over k that can be expressed as the set of the common zeros of a set of homogeneous polynomials $G_j(Y_0, \dots, Y_n)$ is called a **projective algebraic variety** (or simply **projective variety**). In this section, **variety** means either an affine or projective variety. A variety which is a subset of another variety is called a subvariety. These varieties are the forerunners of the modern, more general versions of algebraic varieties, which we will discuss later.

When V is an affine variety in k^n , the set of the polynomials in $k[X] = k[X_1, \dots, X_n]$ that vanish at every point of V form an ideal $I(V)$ of $k[X]$. The residue class ring $A_V = k[X]/I(V)$ is called the **coordinate ring** (or **affine ring**) of V . We can regard A_V as the ring of k -valued functions on V that can be expressed as polynomials of the coordinates of k^n . When V is a projective variety, the \dagger homogeneous ideal generated by the homogeneous polynomials in $k[Y] = k[Y_0, \dots, Y_n]$ that vanish at every point of V is denoted by $I(V)$, and the ring $A_V = k[Y]/I(V)$ is called the **homogeneous coordinate ring** of V .

A variety V is said to be **reducible** or **irreducible** according as it is the union of two proper subvarieties or not. A maximal irreducible subvariety of V is called an **irreducible component** of V . Any variety can be written uniquely as the union of a finite number of irreducible components. A variety V is irreducible if and only if $I(V)$ is a \dagger prime ideal. When that is the case, the field of quotients of A_V (when V is affine) or the subfield of the field of quotients

of A_V , consisting of the homogeneous elements of degree 0 (when V is projective) is called the **function field** of V and is denoted by $k(V)$. Elements of $k(V)$ are called **rational functions** (or simply **functions**) on the variety V . The field $k(V)$ is \dagger finitely generated over k . The transcendence degree of $k(V)$ over k is called the **dimension** of V . When V is reducible, the maximum of the dimensions of its irreducible components is called its dimension. If W is a subvariety of an irreducible variety V , then $\dim V - \dim W$ is called the **codimension** of W on V . A subvariety of pure codimension 1 of an affine or projective space can be defined by a single equation and is called a **hypersurface**. If the ideal $I(V)$ of a variety V of dimension r in a projective space $P^n(k)$ is generated by $n-r$ homogeneous polynomials, then V is called a **complete intersection**. Compared with general varieties, complete intersections have some simpler properties. On the other hand, many important varieties are not complete intersections, e.g., \dagger Abelian varieties of dimension ≥ 2 .

The intersections and finite unions of subvarieties on a variety V are also subvarieties. Thus the subvarieties can be taken as the \dagger system of closed sets of a topology on V (- 426 Topology), which is called the **Zariski topology of the variety V** . When k is the field of complex numbers, V can be viewed as an \dagger analytic space, and the topology of V as such (the "usual" topology) is much stronger than the Zariski topology. For the rest of this article, varieties will be considered as having Zariski topologies unless stated otherwise. Terms such as **Zariski open**, **Zariski closed**, and **Zariski dense** are used to mean open, closed, or dense in a Zariski topology. Suppose a condition (P) concerning the points of an irreducible variety V (concerning the elements of a set M parametrized by the points of V) is satisfied in a nonempty Zariski open set of V . Then we say that the condition (P) holds at **almost all points of the variety V** or at general points of V (almost all elements of the set M).

Let U and V be affine varieties in k^n and k^m , respectively. Then the product set $U \times V$ is an affine variety in k^{n+m} and is called the **product algebraic variety** (or simply the **product**) of U and V . Note that the Zariski topology on $U \times V$ is stronger than the product of the topologies of U and V . When k is \dagger algebraically closed, $U \times V$ is irreducible if U and V are irreducible.

Suppose that k is algebraically closed. Let \mathfrak{P} be a \dagger prime ideal of $k[X] = k[X_1, \dots, X_n]$, and let V be the affine variety in k^n defined as the zero points of \mathfrak{P} . Then $I(V) = \mathfrak{P}$ (\dagger Hilbert zero point theorem) (- 369 Rings of Polynomials D). Therefore there exists a one-to-one correspondence between the set of prime ideals of

$k[X]$ and the set of irreducible varieties in k^n . In particular, the \dagger maximal ideals correspond to the points of k^n . Similarly, there exists a one-to-one correspondence between the set of homogeneous prime ideals of $k[Y]$ other than $\sum_{i=0}^n Y_i k[Y]$ and the set of irreducible subvarieties in $P^n(k)$.

When we deal with nonlinear algebraic equations, we cannot expect a simple, clearcut theory without assuming that k is algebraically closed. Hence we take an algebraically closed field K containing k and regard a variety V in k^n as a subset of the variety V_K in K^n defined by the same equations. From now on, we suppose that k is algebraically closed. If the ideal $I(V)$ of $k[X]$ or $k[Y]$ determined by a variety V is generated by polynomials with coefficients in a subfield k' of k , we say that V is **defined over k'** or that k' is a **field of definition** for V . Any variety has the smallest field of definition, which is finitely generated over the prime field. In the theory of A. Weil [92], we fix an algebraically closed field K that has an infinite transcendence degree over the prime field. This K is called the **universal domain**. A point of V is called a **k' -rational point** of V if all of its coordinates belong to a subfield k' of K . Let K_1, K_2 be two extension fields of a field L , and let $(x) \in K_1^n, (y) \in K_2^n$. We say (y) is a **specialization** of the point (x) over L (notation: $(x) \xrightarrow{L} (y)$) if all polynomials $f(X) \in L[X_1, \dots, X_n]$ satisfying $f(x) = 0$ also satisfy $f(y) = 0$; in other words, if there exists a homomorphism of L -algebras $L[x_1, \dots, x_n] \rightarrow L[y_1, \dots, y_n]$ mapping x_i to y_i . Let K be the universal domain, V an irreducible variety in K^n , and $k' (\subset K)$ a field of definition for V having a finite transcendence degree over the prime field. Then there exists a point (x) of V such that all points of V are specializations of (x) over k' . Such a point (x) (in general not uniquely determined) is called a **generic point** of V over k' . The ring $k'[x]$ is isomorphic to $k'[X]/I(V) \cap k'[X]$ over k' . (Some authors use the term *generic point* to mean *almost all points* as defined earlier.)

B. Local Rings

Let V be an affine variety and let W be an irreducible subvariety of V . Let \mathfrak{P}_W be the subset of A , consisting of the elements that vanish identically on W . Then \mathfrak{P}_W is a prime ideal of A . The ring of \dagger quotients of A , with respect to \mathfrak{P}_W is denoted by $\mathfrak{D}_{V,W}$ or by \mathfrak{D}_W and is called the **local ring** of W on V (or of V at W). Suppose for simplicity that V is irreducible. Then \mathfrak{D}_W is the subring $\{f/g \mid f, g \in A_V, g \notin \mathfrak{P}_V\}$ of $k(V)$, and the \dagger residue field of \mathfrak{D}_W modulo the maximal ideal $\mathfrak{P}_W \mathfrak{D}_W$ can be

identified with $k(W)$. When a function φ on V ($\varphi \in k(V)$) belongs to \mathfrak{D}_W , it is said to be **regular** at W . For a given function $\varphi \in k(V)$, the set of the points of V where φ is regular is Zariski open. In the case of a projective variety, the local ring $\mathfrak{D}_{V,W}$ is defined as the subring of the ring of quotients of A , with respect to \mathfrak{P}_W , consisting of the homogeneous elements of degree 0.

A mapping from an open set U of a variety V to k that is regular at every point of U is called a **regular function** on U . The ring of regular functions on U is denoted by A_U . By assigning A_U to each open set U , we can define a \dagger sheaf of rings \mathcal{O}_V on V , of which the \dagger stalk $\mathcal{O}_{V,x}$ at a point $x \in V$ is the local ring $\mathfrak{D}_{V,x}$. The sheaf \mathcal{O}_V is called the **sheaf of germs of regular functions** on V (or the **structure sheaf** of V) (- 383 Sheaves H).

C. General Definition

Consider a pair (V, \mathcal{O}) of a topological space V and a sheaf \mathcal{O} of germs of mappings from V to k . If V has a finite open covering (U_i) such that each $(U_i, \mathcal{O}|U_i)$ is isomorphic to an affine variety V_i (in the sense that there exists a homeomorphism from U_i to V_i that transforms $\mathcal{O}|U_i$ to the structure sheaf of V_i), the pair (V, \mathcal{O}) is called a **prealgebraic variety over k** , and \mathcal{O} is called its **structure sheaf**. Usually (V, \mathcal{O}) is denoted simply by V .

A **regular mapping** between prealgebraic varieties is defined as a continuous mapping $g: V \rightarrow V'$ satisfying $\varphi \circ g \in \mathcal{O}_{V',x}$ for any $x \in V$ and $\varphi \in \mathcal{O}_{V',g(x)}$. Furthermore, if g is a homeomorphism and g^{-1} is also regular, then g is called an **isomorphism** or a **homeomorphism**. The Cartesian product $X \times Y$ of prealgebraic varieties X and Y is locally a product of affine varieties. Therefore $X \times Y$ has the structure of a prealgebraic variety. A prealgebraic variety X is called an **algebraic variety** if the image of the diagonal mapping $X \rightarrow X \times X$ is closed in the Zariski topology of the product variety $X \times X$ ("separation condition"). (This definition is due to J.-P. Serre [81]). The separation condition corresponds to \dagger Hausdorff's separation axiom. If W is a locally closed subset (i.e., the intersection of an open set and a closed set) of an algebraic variety V , then it becomes an algebraic variety in a natural manner (the germs of regular functions at $P \in W$ are taken to be the germs of functions induced on W by the functions in $\mathcal{O}_{V,P}$). Locally closed subvarieties of k^n or $P^n(k)$ are called **quasi-affine** or **quasiprojective algebraic varieties**, respectively. Definitions of irreducibility and local rings for general algebraic varieties are given in the same manner

as before. In this article, algebraic varieties will often be referred to simply as varieties.

The notion of an irreducible algebraic variety was developed from that of **abstract algebraic variety** (or simply **abstract variety**) defined by Weil.

D. Schemes

The set of prime ideals ($\neq (1)$) of a commutative ring A with unity element 1 is denoted by $\text{Spec}(A)$ and is called the **spectrum** of A . For any subset a of A , we denote by $V(a)$ the set of the prime ideals containing a . We define a topology on $\text{Spec}(A)$ in which the closed sets are $V(a)$. This, again, is called the **Zariski topology** of $\text{Spec}(A)$. For an element f of A , the open set $D(f) = \text{Spec}(A) - V(f)$ is called an elementary open set. The elementary open sets form a base of open sets in the Zariski topology of $\text{Spec}(A)$. The set of closed points is nothing but the set of maximal ideals of A . Assigning to each point \mathfrak{P} of $\text{Spec}(A)$ the \dagger ring of quotients $A_{\mathfrak{P}}$, we obtain a sheaf of rings \tilde{A} on $\text{Spec}(A)$. We have the equality $\Gamma(D(f), \tilde{A}) = A_f$, where A_f is the \dagger ring of quotients by the multiplicative system $\{f^n : n \geq 0\}$. In particular, we have $\Gamma(\text{Spec}(A), \tilde{A}) = A$. Regarded as a \dagger local-ringed space with \tilde{A} as the structure sheaf, $\text{Spec}(A)$ is called an **affine scheme**.

A local-ringed space X which is locally isomorphic to an affine scheme is called a **scheme**. A **morphism of schemes** is, by definition, a \dagger morphism between them as local-ringed spaces. Thus we obtain a \dagger category whose objects are schemes. We denote it by (Sch) . Giving a morphism $f: X \rightarrow \text{Spec}(A)$ is equivalent to giving a ring homomorphism $\Gamma(f): A \rightarrow \Gamma(X, 0)$. Hence the category of affine schemes (which is a \dagger full subcategory of (Sch)) is contravariantly equivalent to the category of commutative rings. If there is given a morphism of schemes $f: X \rightarrow S$, X is said to be an **S-scheme** or a **scheme over S** , and f is called the structure morphism and S the base scheme. For two S -schemes $f: X \rightarrow S$, $g: Y \rightarrow S$, a morphism of S -schemes is defined to be a morphism of schemes $h: X \rightarrow Y$ with $f = g \circ h$. Thus we obtain the category of S -schemes denoted by (Sch/S) . $\text{Spec}(\mathbb{Z})$ is the unique \dagger final object in (Sch) , hence (Sch) is nothing but $(\text{Sch}/\text{Spec}(\mathbb{Z}))$.

The \dagger fiber product always exists in (Sch) . In fact in the case of affine S -schemes $X = \text{Spec}(B)$ and $Y = \text{Spec}(C)$ with $S = \text{Spec}(A)$ we have $X \times_S Y = \text{Spec}(B \otimes_A C)$, and in the general case we construct $X \times_S Y$ by patching together fiber products of affine schemes.

A morphism $f: X \rightarrow S$ is called **separated** if the image of the diagonal morphism $\Delta_{X/S}$:

$X \rightarrow X \times_S X$ is closed. We also say that X is separated over S or X is a **separated S-scheme**. A scheme X is said to be **separated** if it is separated over $\text{Spec}(\mathbb{Z})$. All affine schemes are separated.

When K is a field, $\text{Spec}(K)$ is a space having only one point and equipped with K as the stalk of the structure sheaf. For a point x of a scheme X , denote by $k(x)$ the residue field of $\mathcal{O}_{X,x}$. For $f \in \mathcal{O}_{X,x}$ we call the residue class off in $k(x)$ the value off at x , denoted by $f(x)$. We have a natural morphism $i_x: \text{Spec}(k(x)) \rightarrow X$ whose image is $\{x\}$. More generally, we call a morphism i of a spectrum $\text{Spec}(K)$ of a field K to X a point of X with values in K . Such a point is determined by a point x in X and an embedding of $k(x)$ in K . A point of X with values in an algebraically closed field is called a **geometric point**. For a morphism $f: X \rightarrow S$ and a point s in S , the fiber product $X \times_S \text{Spec}(k(s))$ is called the **fiber** off over s and denoted by $f^{-1}(s)$. For a geometric point $\text{Spec}(K) \rightarrow S$, $X \times_S \text{Spec}(K)$ is called a **geometric fiber**.

A scheme X is called **reduced** if the local ring at each point of X has no \dagger nilpotent elements. A scheme is said to be **irreducible** if its underlying topological space is not a union of two proper closed subsets. A scheme is called **integral** if it is reduced and irreducible. Every local ring of an integral scheme is an \dagger integral domain. If a scheme X has an affine open covering $\{U_i = \text{Spec}(A_i)\}$ such that every A_i is a \dagger Noetherian ring, X is said to be **locally Noetherian**. A locally Noetherian scheme is called **Noetherian** if its underlying topological space is \dagger (quasi-)Compact.

A morphism $f: X \rightarrow Y = \text{Spec}(A)$ is said to be **locally of finite type (of finite type)** if X has an open affine covering (a finite open affine covering) $\{U_i = \text{Spec}(A_i)\}$ such that each A_i is a finitely generated A -algebra. A general morphism $f: X \rightarrow Y$ is said to be **locally of finite type (of finite type)** if there is an open affine covering $\{V_i\}$ of Y such that every restriction of $f: f^{-1}(V_i) \rightarrow V_i$ is locally of finite type (of finite type). If $f: X \rightarrow Y$ is (locally) of finite type we say that X is (locally) of finite type over Y .

A scheme of finite type over a field K (i.e., over $\text{Spec}(K)$) is called an **algebraic scheme** over K . There is a \dagger natural equivalence of categories between the category of reduced separated algebraic K -schemes (as a full subcategory of (Sch/K)) and the category of algebraic varieties over K (defined in Section C) equipped with regular mappings as morphisms. Hence we identify these categories from now on. Occasionally, algebraic variety means irreducible variety. Nonalgebraic schemes are also important as tools for the

study of algebraic varieties. For example, for a point x in a scheme X there is a canonical monomorphism $j_x: \text{Spec}(\mathcal{O}_{X,x}) \rightarrow X$ by which the unique closed point of $\text{Spec}(\mathcal{O}_{X,x})$ is mapped to x . If for two algebraic K -schemes X, Y and for two points $x \in X, y \in Y$ there is a K -isomorphism $\mathcal{O}_{X,x} \cong \mathcal{O}_{Y,y}$, then suitable neighborhoods of x and y are isomorphic over K .

Many concepts concerning varieties, e.g., dimension, generic points, specialization, can be naturally extended to the case of schemes by virtue of commutative ring theory.

A morphism of schemes $f: X \rightarrow Y$ is called **proper** if it satisfies the following two conditions: (1) f is separated and of finite type, (2) for every scheme T and for every morphism $T \rightarrow Y$, the morphism $X \times_Y T \rightarrow T$ obtained from f by the "change of base" is a closed mapping. We also say that X is a proper Y -scheme or X is proper over Y . A proper algebraic K -scheme is called **complete**. A projective variety is complete, while an affine variety over K is complete only when it is of dimension zero. Every algebraic variety can be embedded in a complete variety (M. Nagata [63]).

A morphism of schemes $f: X \rightarrow Y$ is called **affine** if every inverse image by f of an open affine subset of Y is again an affine scheme.

A morphism of schemes $f: X \rightarrow Y$ is called **finite** if it is of finite type and there is an affine open covering $\{U_i = \text{Spec}(A_i)\}$ of Y such that $f^{-1}(U_i) = \text{Spec}(B_i)$, where B_i is integral over A_i . For a locally Noetherian scheme Y and a morphism of schemes $f: X \rightarrow Y$ the following three conditions are equivalent: (i) f is finite; (ii) f is affine and proper; (iii) f is proper and every fiber off is a finite set. For a finite surjective morphism of Noetherian schemes $f: X \rightarrow Y$, X is an affine scheme if and only if Y is an affine scheme.

A morphism of schemes $f: X \rightarrow Y$ is said to be **flat** if for each point $x \in X$, $\mathcal{O}_{X,x}$ is a flat $\mathcal{O}_{Y,f(x)}$ -module. If, moreover, f is surjective, then f is called **faithfully flat**. Assume that $g: Y' \rightarrow Y$ is a faithfully flat morphism of finite type of locally Noetherian schemes and $f: X \rightarrow Y$ is a morphism of schemes. Then for many important properties of morphisms, f has these properties if and only if the pull-back $f_{Y'}: X \times_Y Y' \rightarrow Y'$ has the same properties (theory of descent [29, 30]).

E. Cohomology Theory

Let (X, \mathcal{O}) be a ringed space. An \mathcal{O} -Module (i.e., a sheaf of \mathcal{O} -modules) F is said to be **quasicoherent** if for each point x of X there exist a neighborhood U of x and an exact sequence $M \rightarrow N \rightarrow F|_U \rightarrow 0$, where M and N are free \mathcal{O}_U -Modules. An \mathcal{O} -Module F is said to

be **of finite type** if F is locally generated by a finite number of sections over \mathcal{O} ; F is **of finite presentation** if, locally, there exists an exact sequence $\mathcal{O}^p \rightarrow \mathcal{O}^q \rightarrow F \rightarrow 0$ where p and q are positive integers (they need not be globally constant); F is **coherent** if (i) F is of finite type and (ii) the kernel of any homomorphism $\mathcal{O}_U^n \rightarrow F|_U$ (where n is an arbitrary positive integer, and U is an open set) is of finite type. Obviously, if F is coherent, then F is of finite presentation, which implies that F is quasicoherent and of finite type. In the category of \mathcal{O} -Modules, the full subcategory of coherent sheaves is closed under almost all operations of sheaves. If \mathcal{O} itself is coherent as an \mathcal{O} -Module, \mathcal{O} is said to be a **coherent sheaf of rings**. In this case every \mathcal{O} -Module of finite presentation is coherent.

The structure sheaf of a locally Noetherian scheme is a coherent sheaf of rings. On a locally Noetherian scheme X , every quasicoherent sub- \mathcal{O}_X -Module or quotient \mathcal{O}_X -Module of a coherent \mathcal{O}_X -Module is coherent. A coherent \mathcal{O}_V -Module on an algebraic variety V is called a **coherent algebraic sheaf**.

Let $X = \text{Spec}(A)$ be an affine scheme. Then every quasicoherent \mathcal{O}_X -Module F on X is generated by its global sections. The correspondence $F \rightarrow \Gamma(X, F)$ defines an equivalence between the category of quasicoherent sheaves on X and the category of A -modules; if A is Noetherian, then the coherent sheaves and the finite A -modules correspond to each other under this equivalence.

Let X be a separated scheme, and $\mathcal{U} = \{U_i\}$ an affine open covering of X . For each quasicoherent \mathcal{O}_X -Module F , the cohomology group $H^q(X, F)$ is canonically isomorphic to the Čech cohomology $H^q(\mathcal{U}, F)$ (- 383 Sheaves F). If X is of dimension d , then $H^q(X, F) = 0$ for every sheaf F of Abelian groups on X and $q > d$.

For a scheme X we define the **cohomological dimension** $\text{cd}(X)$ to be the largest integer q such that $H^q(X, F) \neq 0$ for a quasicoherent \mathcal{O}_X -Module F on X [35]. The cohomological dimension $\text{cd}(X)$ does not exceed the dimension of X . If X is an affine scheme, then $\text{cd}(X) = 0$. The converse is true under the assumption that X is Noetherian (Serre's criterion [29, III]). For an algebraic scheme X of dimension n , $\text{cd}(X) = n$ if and only if X is complete (S. L. Kleiman).

Let $f: X \rightarrow Y$ be a proper morphism of Noetherian schemes. Then for every coherent \mathcal{O}_X -Module F and for every $q \geq 0$, $R^q f_*(F)$ (- 200 Homological Algebra) is also coherent. In the special case of $Y = \text{Spec}(k)$ with a field k this means that for an algebraic coherent sheaf F on a complete variety X the cohomology group $H^q(X, F)$ is a finite-dimensional vector space over k .

Let X be a scheme over k and let F be a locally free \mathcal{O}_X -Module of rank n (i.e., an \mathcal{O}_X -Module which is locally isomorphic to \mathcal{O}_X^n). If we take an open covering $\{U_i\}$ of X and isomorphisms $\varphi_i: F|_{U_i} \xrightarrow{\sim} \mathcal{O}_{X|U_i}^n$, then $\varphi_i \circ \varphi_j^{-1}$ defines a morphism $g_{ij}: U_i \cap U_j \rightarrow GL(n, k)$, which is called the coordinate transformation of F . If we construct a \dagger vector bundle B on X by the same coordinate transformations g_{ij} , then F can be regarded as the sheaf $\mathbf{B}(B)$ of germs of sections of B . By means of the canonical homomorphism $GL(n, k) \rightarrow PGL(n-1, k)$ we can construct a projective bundle $\mathbf{P}(F)$ on X (which is said to be associated with F). (Note that in [29], $\mathbf{P}(F)$ is defined to be a projective bundle with coordinate transformations ${}^t g_{ij}^{-1}$, i.e., associated with the dual of F in our sense.) This procedure of associating $\mathbf{P}(F)$ with a locally free \mathcal{O}_X -Module F can be generalized for any quasicoherent \mathcal{O}_S -Module of finite type on an arbitrary scheme S .

A closed (locally closed) S -subscheme $f: X \rightarrow S$ of $p: \mathbf{P}(F) \rightarrow S$ is called a **projective scheme** (quasiprojective scheme) **over** S , or f is said to be a **projective morphism** (quasiprojective morphism). A projective morphism is proper. A reduced projective scheme over a field k is nothing but a projective variety over k . We can develop the theory of projective schemes by means of \dagger graded rings in a way similar to affine schemes.

A locally free \mathcal{O}_X -Module of rank 1 is called an **invertible sheaf**. Invertible sheaves correspond to line bundles up to isomorphisms. Let $P = \mathbf{P}^N(k)$ be a projective space, (y_0, y_1, \dots, y_N) a system of homogeneous coordinates of P , and U_i the open subset of P defined by $y_i \neq 0$. Denote by $\mathcal{O}(n)$ the invertible sheaf on P defined by the coordinate transformation $g_{ij} = (y_j/y_i)^n$. More generally, let $p: P = \mathbf{P}(F) \rightarrow S$ be the projective bundle associated with a locally free \mathcal{O}_S -Module F of rank $N+1$ on a scheme S . Then there is an invertible sheaf $\mathcal{O}(n) = \mathcal{O}(1)^{\otimes n}$ on P with the properties: (i) for each $s \in S$ its restriction to the fiber $p^{-1}(s) = \mathbf{P}^N(k(s))$ is $\mathcal{O}(n)$ defined above; (ii) $p_*(\mathcal{O}(n)) = S^n(F)$ for $n > 0$ where $S^n(F)$ denotes the n th symmetric product of F . The invertible sheaf $\mathcal{O}(1)$ is called the **tautological line bundle** on P . (Note that $\mathcal{O}(1)$ in the sense of [29] is $\mathcal{O}(-1)$ in our sense, but since the definition of $\mathbf{P}(F)$ is also different, the above property (ii) holds without modification.)

For a quasiprojective S -scheme $f: X \rightarrow \mathbf{P}(E) \rightarrow S$, the restriction of $\mathcal{O}(1)$ to X is denoted by $\mathcal{O}_X(1)$ (or simply $\mathcal{O}(1)$). An invertible sheaf L on X is called **very ample over** S if there exist a locally free \mathcal{O}_S -Module of finite type E on S and an S -immersion $i: X \rightarrow \mathbf{P}(E)$ such that $\mathcal{O}_X(1) = L$; L is called **relatively ample** over S or S -ample if $L^{\otimes n}$ is very ample over S for a cer-

tain $n > 0$ (- Section N). When S is an affine scheme, an ample (very ample) sheaf over S is simply called **ample** (**very ample**). There is the following cohomological criterion of ampleness (generalized **Serre's theorem**).

Let Y be a Noetherian scheme, $f: X \rightarrow Y$ a proper morphism, and L an invertible \mathcal{O}_X -Module. Then the following four conditions are equivalent: (i) L is f -ample; (ii) for each coherent \mathcal{O}_X -Module F there is an integer N such that $R^q f_*(F \otimes L^{\otimes n}) = 0$ for all $n \geq N$ and $q > 0$; (iii) for each coherent sheaf of ideals \mathcal{I} of \mathcal{O}_X there is an integer N such that $R^1 f_*(\mathcal{I} \otimes L^{\otimes n}) = 0$ for all $n \geq N$. They imply the condition (iv): for each coherent \mathcal{O}_X -Module F there is an integer N such that the canonical homomorphisms $f_* f_*(F \otimes L^{\otimes n}) \rightarrow F \otimes L^{\otimes n}$ are surjective for all $n \geq N$.

Let X be a scheme proper over a field k (such an X is called a **k -complete scheme**), and let F be a coherent \mathcal{O}_X -Module. Then all $H^q(X, F)$ are finite-dimensional k -vector spaces (finiteness theorem), and $\chi(F) = \sum (-1)^q \dim H^q(X, F)$ is a finite number, called the **Euler-Poincaré characteristic** of F over X . For every invertible sheaf L , $\chi(F \otimes L^{\otimes m})$ is a polynomial in m , which is said to be the **Snapper polynomial** [48]. $\text{Supp } F$, defined to be the set $\{x \in X \mid F_x \neq 0\}$, is a closed subset of X and is called the **support** of F . The degree of the polynomial $\chi(F \otimes L^{\otimes m})$ is at most $\gamma = \dim \text{supp } F$, and we have $\chi(F \otimes L^{\otimes m}) = e \cdot m^\gamma / \gamma! + \text{lower terms in } m$ for some $e \in \mathbf{Z}$. e is said to be the **intersection number** of L with F [48]. When W is a closed subscheme of X defined by an \mathcal{O}_X -ideal I , $(L^\cdot \cdot W)$ is defined to be the intersection number of L with \mathcal{O}_X/I , where $r = \dim W$. If L is ample, then $(L^\cdot \cdot W) > 0$ for any W . In particular, when $W = X$, it follows that $(L^n) > 0$, where $n = \dim X$. The converse of this fact is the **Nakai-Moishezon criterion**, saying that if $(L^\cdot \cdot W) > 0$ for any closed subvariety W of X where $r = \dim W$, then L is ample [56, 64].

When L is ample, $H^q(X, F \otimes L^{\otimes m}) = 0$ for any $q > 0$ and for sufficiently large m ; hence $\chi(F \otimes L^{\otimes m}) = \dim H^0(X, F \otimes L^{\otimes m})$ and the Snapper polynomial $\chi(F \otimes L^{\otimes m})$ is also called the **Hilbert polynomial** or the **Hilbert characteristic function** of F . If L is the invertible sheaf defined by a hyperplane section of X in \mathbf{P}_k^N , i.e., $L = \mathcal{O}_X(1)$, then $\chi(L^{\otimes m}) = \dim H^0(X, L^{\otimes m}) = \dim R_m$ for sufficiently large m where $\bigoplus R_m$ denotes the homogeneous coordinate ring of X in \mathbf{P}_k^N .

In general, for a complete irreducible variety V of dimension γ , we put

$$\chi(V) = \chi(\mathcal{O}_V) = \sum_{q=0}^{\gamma} (-1)^q h^{0,q}$$

($h^{0,q} = \dim H^q(V, \mathcal{O}_V)$) and call it the **arithmetic**

genus of V . Classically, the number $p_a(V)$ defined by $p_a(V) = (-1)^r(\chi(V) - 1) = h^{0,r} - h^{0,r-1} + \dots \pm h^{0,1}$ was called the arithmetic genus of V , instead of $\chi(V)$. When V is a nonsingular irreducible curve, $p_a(V)$ is the usual \dagger genus. If V is a projective variety, the constant term of the Hilbert polynomial is $\chi(V)$ and the coefficient of its highest term is $(\deg V)/r!$.

Let V be a \dagger normal variety, and D a \dagger divisor on V . If, for each point $x \in V$, we denote by $\mathcal{O}(D)_x$ the set of the functions $f \in k(V)$ that satisfy $(f) + D \geq 0$ on some neighborhood of x , we obtain a coherent algebraic sheaf $\mathcal{O}(D)$, and we have $\dim H^0(V, \mathcal{O}(D)) = I(D)$. If, moreover, V is complete, we put $\chi_V(D) = \chi(V) - \chi(\mathcal{O}(-D))$ and call it the **virtual arithmetic genus** of D . Classically, the number $p(D) = (-1)^{r-1}(\chi_V(D) - 1)$ was called by that name. When D is effective and has no multiple components, $\chi_V(D)$ coincides with the arithmetic genus of D regarded as a variety. In general, $\chi_V(D)$ stays invariant if D is replaced by a divisor that is algebraically equivalent to D (\rightarrow Section N).

If D is a Cartier divisor, $\mathcal{O}(D)$ is an invertible sheaf. For two Cartier divisors D_1, D_2 , $\mathcal{O}(D_1 + D_2) = \mathcal{O}(D_1) \otimes \mathcal{O}(D_2)$, and D_1 and D_2 are linearly equivalent if and only if $\mathcal{O}(D_1) \cong \mathcal{O}(D_2)$.

Let V be an irreducible nonsingular variety and let Ω^p denote the sheaf of germs of regular differential forms of degree p ($\Omega^0 = \mathcal{O}_V$). If V is complete, then we denote $\dim H^q(V, \Omega^p)$ by $h^{p,q}$.

Serre's duality theorem: Let V be a nonsingular complete irreducible variety of dimension r , B an algebraic vector bundle over V , and B^* the dual vector bundle of B . Denote by \mathcal{B} and \mathcal{B}^* the sheaves of germs of sections of B and B^* , respectively. Then (i) $H^r(V, \Omega^r)$ is canonically isomorphic to k , and (ii) $H^q(V, \mathcal{B})$ and $H^{r-q}(V, \mathcal{B}^* \otimes \Omega^r)$ are dual to each other as linear spaces by means of the \dagger cup product of the above spaces with $H^r(V, \Omega^r) \cong k$. In particular, $H^q(V, \Omega^p)$ is dual to $H^{r-q}(V, \Omega^{r-p})$; hence we have $h^{p,q} = h^{r-p, r-q}$.

This theorem was extensively generalized by Grothendieck in the category of schemes [1, 331].

When the field k is of characteristic 0, we furthermore have $h^{p,q} = h^{q,p}$ by complex conjugation (\rightarrow 232 Kähler Manifolds C); but in characteristic p , there are examples for which this symmetry does not hold [61]. In general, $h^{p,q}$ is a \dagger relative invariant but not an \dagger absolute invariant; however, as $h^{p,0}$ is the dimension of the linear space of the differential forms of the first kind with degree p , $h^{p,0}$ is an absolute invariant. Hence $h^{0,p}$ is also an absolute invariant in the case of characteristic 0. When the characteristic is positive, the absolute invariance of $h^{0,p}$ has not yet been proved.

F. Simple Points and Singular Points

Let V be a variety over an algebraically closed field k . We say that a point P of V is **simple** or that V is **nonsingular** or **smooth** at P if the local ring \mathfrak{O}_P is a \dagger regular local ring. Since the problem is local, we may assume that V is an affine variety in k^n . Then the simplicity of P on V is equivalent to the following condition: P is contained in only one irreducible component of V , and if that component has dimension r there exist $n-r$ polynomials $F_i(X)$ in $I(V)$ such that $\text{rank}(\partial F_i / \partial X_j)_{(X)=P} = n-r$. A point of V that is not simple is called a **singular point** or a **multiple point**. The set of singular points on V (called the **singular locus** of V) is a proper closed subset of V . A variety with no multiple points is called **smooth** or **nonsingular**.

This notion can be made relative. A morphism $f: X \rightarrow Y$ of a locally Noetherian scheme is called **smooth** if f is flat and locally of finite type and all the geometric fibers off are nonsingular. In the case of an affine morphism $f: X = \text{Spec}(R[X_1, \dots, X_{r+s}]/(f_1, \dots, f_m)) \rightarrow Y = \text{Spec}(R)$ of relative dimension r (by which we mean the dimension of the general fiber) with a Noetherian ring R , the smoothness off amounts to a condition that $\text{rank}((\partial f_i / \partial X_j)(x)) = s$ at each point x of X .

When for a point P of a variety V the local ring \mathfrak{O}_P is \dagger normal, P is called a **normal point**. A simple point is normal. The set of normal points is a nonempty open subset of V . An irreducible variety whose points are all normal is called a **normal algebraic variety** (or simply **normal variety**). The singular locus of a normal variety has codimension ≥ 2 . For an irreducible variety V , there exists a pair (V', f) of a normal variety V' and a birational finite morphism $f: V' \rightarrow V$; V' is unique up to isomorphisms and is called the **derived normal model** or **normalization** of V .

Simplicity and normality for V at a subvariety W are defined in the same way as at a point by using the local ring $\mathfrak{O}_{V,W}$.

For a morphism $f: X \rightarrow Y$ of locally Noetherian schemes, locally of finite type, the following three conditions are equivalent: (i) f is smooth and every fiber off is a discrete set; (ii) f is flat and every geometric fiber over $\text{Spec}(K)$ off is a union of spectra of fields isomorphic to K ; (iii) f is flat and every fiber off over $y \in Y$ is a union of spectra of fields that are finite \dagger separable extensions of $k(y)$. These conditions are local with respect to X . If a morphism f satisfies these equivalent conditions, we say that f is **étale** or X is **étale** over Y . A morphism

$$f: X = \text{Spec}(R[X_1, \dots, X_n]/(f_1, \dots, f_n)) \rightarrow Y = \text{Spec}(R)$$

is étale if and only if $\det((\partial f_i / \partial X_j)(x)) \neq 0$ for all $x \in X$. Hence étale morphisms correspond to local isomorphisms in the analytic category. For a surjective étale morphism $f: X \rightarrow Y$ many important geometric properties (reduced, integral, normal, nonsingular, etc.) hold on X if and only if they hold on Y (theory of descent).

G. Dimension Theorems

Let V be an irreducible variety, and let U and W be irreducible subvarieties of V . Then any irreducible component of $U \cap W$ that is simple on V has dimension $\geq \dim U + \dim W - \dim V$. When the equality holds, the component is called a **proper component** of the intersection $U \cap W$. If every component of $U \cap W$ that is simple on V is proper, we say that U and W **properly intersect** on V . Any two subvarieties U and W of $\mathbb{P}^n(k)$ with $\dim U + \dim W \geq n$ intersect each other. When V is an irreducible r -dimensional variety in $\mathbb{P}^n(k)$, the number of points of intersection $V \cap L$ of V with an $(n-r)$ -dimensional linear variety L is independent of the choice of L as long as L is in a "general position." This number is called the **degree** of V and is denoted by $\deg(V)$. Letting $\mathcal{O}(1)$ be the fundamental sheaf of $\mathbb{P}^n(k)$, we have $\deg V = (\mathcal{O}(1)^r \cdot V)$ (- Section E).

H. Group Varieties

An algebraic variety G is called an **algebraic group** if it has a group structure and if the mapping $G \times G \rightarrow G$ sending (x, y) to xy^{-1} is a morphism. Every algebraic group is quasi-projective (Chow). If G is irreducible, then it is also called a **group variety**; a complete group variety is called an **Abelian variety** (- 3 Abelian Varieties B, 13 Algebraic Groups B). A scheme G over another scheme S equipped with morphisms over $S: G \times_S G \rightarrow G, G \rightarrow G$, and $S \rightarrow G$, called multiplication, inverse, and unit section, respectively, which satisfy the relations corresponding to the usual axioms of group, is called a **group scheme** (over S). As a point set, G is not a group, while, for any scheme T over S the set $G(T) = \text{Hom}_S(T, G)$ of the morphisms from T to G is a group (- 52 Categories and Functors M). Consider an algebraic group scheme G over $S = \text{Spec}(k)$. If the characteristic of k is zero, then G is necessarily reduced, so an algebraic group scheme over k is essentially the same as an algebraic group; if k has characteristic p , there exist algebraic group schemes over k that are not reduced.

I. Rational Mappings

Let $f: V \rightarrow V'$ be a morphism of varieties. If V is not complete, the image $f(V)$ is not always closed; the closure $\text{off}(V)$ (in V') is called the **closed image** of V . The image $f(V)$ contains an open dense subset of the closed image.

Let V and W be irreducible varieties. A closed subset T of $V \times W$ is called an **algebraic correspondence** of V and W . We say that points $P \in V$ and $Q \in W$ correspond to each other by T if $(P, Q) \in T$. If T is irreducible and the closed image of the projection $T \rightarrow V$ coincides with V , then the function field $k(V)$ can be identified with a subfield of $k(T)$; if we have $k(V) = k(T)$ with this identification, then T is called a **rational mapping** from V to W . Moreover, if the same conditions are satisfied for W , then T is called a **birational mapping** (of **birational correspondence** or **birational transformation**), and in this case we have $k(V) = k(W)$. A morphism can be considered a special kind of rational mapping by taking the graph. If T is a rational mapping from V to W and W_1 is the closed image of T in W , then $k(W_1)$ can be regarded as a subfield of $k(T) = k(V)$. If $k(V)$ is \dagger separably generated (\dagger purely inseparable) over $k(W_1)$, then T is said to be **separable** (purely inseparable).

Let T be a rational mapping from V to W , and let V' and W' be irreducible subvarieties of V and W , respectively. If there exists an irreducible subvariety T' of T whose projections have the closed images V' and W' , then we say V' and W' correspond to each other by T . The union of irreducible subvarieties of W that correspond to V' by T is a closed subset of W ; it is called the **proper transform** of V' by T and is denoted by $T[V']$. Note that $V' \supset V''$ does not imply $T[V'] \supset T[V'']$. The set of points of W that correspond to the points of V' is called the **total transform** of V' by T and is denoted by $T\{V'\}$. Identifying $k(T)$ with $k(V)$, we have $\mathfrak{O}_{T, T'} \supset \mathfrak{O}_{V', V'}$ in general; if the equality holds, we say that T is **regular** (or **defined**) **along** V' . In that case, W' is the unique irreducible subvariety of W corresponding to V' by T . If V' is simple and of codimension 1 in V , then T is always regular along V' . The set U of points of V at which T is regular is a nonempty open subset, and the restriction of T to U defines a morphism from U to W . A rational mapping can be defined as the closure of the graph of a morphism **defined** on an open subset of V . The study of a rational mapping can be reduced to that of the birational morphism $T \rightarrow V$. **Zariski's main theorem:** Let $S: X \rightarrow Y$ be a birational mapping, and assume that the inverse $S^{-1}: Y \rightarrow X$ is regular, and that X is normal along an irreducible subvariety X' . If there exists an irre-

ducible component Y' of $S[X']$ with $\dim X' \geq \dim Y'$, then S is regular along X' . It follows from the above main theorem that, if $T: V \rightarrow W$ is a rational mapping and if P is a normal point of V such that $T[P]$ contains an isolated point, then T is regular at P .

For a birational mapping $T: V \rightarrow W$ between complete irreducible varieties, a subvariety V' of V is said to be **fundamental** when $\dim T[V'] > \dim V'$. When V' is a point (curve) we say V' is a **fundamental point** (**fundamental curve**) with respect to T . The most classical example of a birational correspondence with fundamental points is the **quadratic transformation** T of a projective plane onto itself given by $(x_0: x_1: x_2) \rightarrow (x_1x_2: x_2x_0: x_0x_1)$ with a suitable coordinate system. In this case, $T^2 = \text{identity}$ and the points P_i defined by $x_j = 0$ for $j \neq i$ correspond to the lines $x_i = 0$ by T . Let Cr denote the group of birational transformations of $\mathbf{P}^n(\mathbf{k})$ into itself. These transformations are called **Cremona transformations** when $n \geq 2$. Cr is generated by linear transformations and quadratic transformations (M. Noether). Recently, Cr has been studied in detail by M. Demazure [21] and H. Umemura [91]. Let V be a complete nonsingular irreducible variety over k . It is called **relatively minimal** if every birational morphism from V to a complete nonsingular variety V' is an isomorphism. It is called **minimal** if every birational mapping from a nonsingular variety V' to V is a morphism. Replacing birational mapping by rational mapping, we can define **strong (absolute) minimality**. Abelian varieties and complete nonsingular curves with positive genera are strongly minimal. In general, minimality implies relative minimality.

Zariski's main theorem is closely related to the **general connectedness theorem** due to W. Fulton and J. Hansen stated as follows: Let P be the product of r copies of \mathbf{P}_k^m and A be the diagonal subvariety of P . If X is an irreducible projective variety and $f: X \rightarrow P$ is a morphism with $\dim f(X) > (r-1)m$, then $f^{-1}(\Delta)$ is connected [22]. The following results are derived from this theorem: (1) if X is singular with only normal crossings (- Section L), then X cannot be imbedded in \mathbf{P}^{2n-1} , where $n = \dim X$. (2) Let X be a nonsingular subvariety of \mathbf{P}^N which is not contained in any hyperplane. Suppose that by the projection $\pi_x: \mathbf{P}^N \rightarrow \mathbf{P}^{N-1}$ with center x , x being a general point of \mathbf{P}^N , X is isomorphic to $\pi_x(X)$. Then $3 \cdot \dim X \leq 2(N-2)$ (Zak's theorem). (3) Let X be a nonsingular subvariety of \mathbf{P}^N and H be an arbitrary hyperplane section of X . Then $\dim(\text{Sing } H) < \text{codim } X$ (J. Roberts).

By a similar connectedness theorem, it was proved that if C is a (reducible) curve with normal crossings on \mathbf{P}^2 over C , then the fun-

damental group of $\mathbf{P}^2 - C$ is commutative, which was conjectured by O. Zariski [20].

J. Rational Varieties

An irreducible algebraic variety V over k whose function field is purely transcendental over k is called a **rational variety**. A complete smooth surface S over an algebraically closed field is rational if and only if $P_2(S) = q(S) = 0$ (the Castelnuovo-Zariski criterion, \rightarrow 15 Algebraic Surfaces E).

If the function field of V has a finite algebraic extension which is purely transcendental over k , then V is called **unirational**. A unirational curve is in fact rational. More generally, if the function field $k(C)$ of a curve C over k is contained in a field finitely generated and purely transcendental over k , then C is rational (**Lüroth's theorem**). A unirational surface over an algebraically closed field of characteristic zero is rational by virtue of the above criterion, but in the case of positive characteristic there are unirational surfaces which are not rational (Zariski). There are nonrational unirational threefolds even of characteristic zero; for example, all smooth cubic hypersurfaces in \mathbf{P}^4 (C. H. Clemens and P. A. Griffiths [16]), and some smooth quartic hypersurfaces in \mathbf{P}^4 (V. A. Iskovskii and Yu. I. Manin [45]). See [6].

K. Monoidal Transformations

Let V be an irreducible variety and \mathcal{J} be a sheaf of ideals of \mathcal{O}_V . For any affine open set U of V , $\mathcal{J}|U$ is determined by an ideal a of the coordinate ring A of U . Let a, a_1, \dots, a_m be a system of generators of a , and let U' be the graph of a rational mapping from U to $\mathbf{P}^m(k)$ such that the points $P \in U$ and $(a_0(P): a_1(P): \dots : a_m(P)) \in \mathbf{P}^m(k)$ correspond to each other by U' . Then U' is uniquely determined (up to isomorphisms) by U and a only. Suppose V has a covering by affine open sets U_i . We obtain U'_i over U_i as before. By patching them together, we get a birational morphism $T: V' \rightarrow V$, which is unique up to isomorphisms. This T with V' is called the **monoidal transformation** or **blowing up** of V by the ideal sheaf \mathcal{J} . Note that in some references, the inverse transformation T^{-1} is said to be the monoidal transformation. The inverse image ideal $T^{-1}(\mathcal{J})$ is an invertible sheaf, that is relatively ample over V . If W is the support of $\mathcal{O}_V/\mathcal{J}$ and $W \neq V$, then $T^{-1}(W)$ has codimension 1 and T gives rise to the isomorphism from $V' - T^{-1}(W)$ onto $V - W$. Thus one can say that V' is obtained from V by replacing W with $T^{-1}(W)$, which is locally defined by prin-

principal ideals. If every point of W is a nonsingular point of V and W is itself nonsingular, $T^{-1}(W)$ is the projective bundle over W whose fiber is the projective r -space where $r = \dim V - \dim W - 1$. In general, if W is a subvariety defined by a sheaf of ideals \mathcal{J} , then T is said to be the **monoidal transformation with center W** . In particular, if W is a point, T is called the (locally) **quadratic transformation**.

L. Resolution of Singularities

Given an arbitrary irreducible variety V , we have the problem of finding out a nonsingular projective variety V' birationally equivalent to V . This is called the problem of **resolution of singularities**. In the case of characteristic zero, this problem was solved by Zariski (1944) for dimension ≤ 3 and by H. Hironaka (1964) [37] for any dimension. In the case of characteristic p , S. Abhyankar solved the 2-dimensional case (1956) and the 3-dimensional case (1966). Hironaka's theorem of resolution of singularities is stated as follows: Let V be a variety over a field of characteristic zero. Then there exists a finite sequence of morphisms of varieties: $V_r \rightarrow V_{r-1} \rightarrow \dots \rightarrow V_1 \rightarrow V_0 = V$ such that (1) V_r has no singular points, (2) each $V_{i+1} \rightarrow V_i$ is a monoidal transformation of V_i with center D_i , (3) each D_i is a nonsingular subvariety, (4) each V_i is normally flat along D_i . Here, V is said to be **normally flat** along a subscheme D of V defined by the sheaf of ideals \mathcal{J} , if the quotient modules $\mathcal{J}_x^p / \mathcal{J}_x^{p+1}$ are flat $\mathcal{O}_{D,x}$ -modules for all p and all points x of D .

Let V be a nonsingular variety and D an effective divisor on V . D is said to be a divisor with (only) **normal crossings** at $x \in V$, if D is defined by $f_1 \cdot f_s$ such that (f_1, \dots, f_s) is a part of a system of local coordinates around x . D is said to be a divisor with only normal crossings, if it is so everywhere. For any subvariety W on V , there exist $h: \tilde{V} \rightarrow V$ that is a composition of monoidal transformations with nonsingular centers, such that $h^{-1}(W)$ has only normal crossings. This results from Hironaka's main theorem II [37]. The normal crossing divisor is also defined for complex manifolds, and similar results hold.

M. Cycles and Divisors

Let V be an irreducible variety. We denote by \mathfrak{B}_r the set of r -dimensional irreducible subvarieties of V that are simple on V (i.e., are not contained in the singular locus of V), and by $\mathfrak{Z}_r(V)$ the free Abelian group with basis \mathfrak{B}_r . Elements of $\mathfrak{Z}_r(V)$ are called cycles of dimension r (or r -cycles) on V . Let A and B be r -

cycles; $A = \sum n_i A_i, B = \sum m_i A_i (A_i \in \mathfrak{B}_r, A_i \neq A_j \text{ if } i \neq j)$. If $n_i \geq m_i$ for all i , then we write $A \geq B$. If $A \geq 0$, then A is said to be a **positive cycle**. For a 0-cycle $A = \sum n_i P_i$ the integer $\deg(A) = \sum n_i$ is called the **degree** of A .

A cycle of codimension 1 is called a **divisor**. A divisor ≥ 0 is usually called **effective** instead of positive. If V is of dimension d and if $W \in \mathfrak{B}_{d-1}$, the local ring $\mathcal{O}_{V,W}$ is a discrete valuation ring. The normalized valuation defined by it is denoted by $v_W(\cdot)$. For a function $f \in k(V)$, we say that W is a **zero** of order n if $v_W(f) = n > 0$, and that W is a **pole** of order $-n$ if $v_W(f) = n < 0$. Any function $f \in k(V)$, other than the constant 0, has at most a finite number of zeros and poles. We denote by $(f)_0$ the sum $\sum v_W(f) W$ extended over all the zeros of f , and put $(f^{-1})_0 = (f)_\infty$ and $(f)_0 - (f)_\infty = (f)$. We call $(f)_0$, $(f)_\infty$, and (f) the **zero divisor**, the **pole divisor**, and the **divisor** of f , respectively. The divisor (f) is equal to $\sum v_W(f) W$, where the summation is taken over $W \in \mathfrak{B}_{d-1}$, and we have $(fg) = (f) + (g)$. When V is complete and the singular locus of V has codimension > 1 , then (f) is constant if and only if $(f) = 0$ (or $(f)_0 = 0$). Let D_1 and D_2 be divisors; if there exists a function $f (\neq 0) \in k(V)$ such that $D_1 - D_2 = (f)$, then D_1 and D_2 are called **linearly equivalent** to each other and we write $D_1 \sim D_2$. The **linear equivalence class** containing a divisor D is denoted by $\text{cl}(D)$. A divisor which is linearly equivalent to 0 on a neighborhood of each point of V is called a **Cartier divisor** (some authors call a Cartier divisor simply a divisor). If V is smooth, then any divisor is a Cartier divisor. If a divisor D can be written as $D = (f)$ on an open set U , then the function f is called a **local equation** of D on U . Let $T: V' \rightarrow V$ be a rational mapping from a normal variety V' to a complete variety V , let D be a Cartier divisor on V , and assume that the closed image of T is not contained in D . Since T is regular on some open set U such that $\text{codim}(V' - U) > 1$, we have a morphism $\varphi = T|_U$ and the pullback $\varphi^*(D)$ defined by composing the local equations of D with φ . Taking the closure of this divisor in V' , we obtain a divisor on V' , which is denoted by $T^*(D)$.

N. Divisors and Linear Systems

Let V be a complete irreducible variety, f_0, f_1, \dots, f_n be elements of the function field $k(V)$ of V , and D be a divisor on V satisfying (fi) $+ D \geq 0$ for each i . Then the set Σ of the divisors of the form $(\sum a_i f_i) + D$, where the a_i are elements of k and not all zero, is called a **linear system**. The linear space $kf_0 + kf_1 + \dots + kf_n$ is called a **defining module** of Σ . The divisors in

Σ are positive and are linearly equivalent to each other; if every positive divisor that is linearly equivalent to a member of Σ belongs to Σ , then Σ is said to be a **complete linear system**. For any linear system Σ , there exists a unique complete linear system containing it, which is denoted by $|\Sigma|$. The maximal positive divisor D_0 that is contained in all divisors of Σ is called the **fixed component** of Σ , and for each DEC we call $D - D_0$ the **variable component** of D (or of Σ). A point P of V is called a **base point** of a linear system Σ if P is on each variable component of Σ . A linear system Σ is called **irreducible** if its generic member is irreducible; otherwise it is called **reducible**. The dimension of a defining module of a linear system Σ is denoted by $l(C)$; we call $l(C) - 1$ the **dimension** of Σ and denote it by $\dim \Sigma$. A linear system of dimension 1 is called a **linear pencil**.

A defining module of a linear system Σ is determined uniquely up to k -isomorphisms. Let L be a defining module, and let f_0, f_1, \dots, f_n be a linearly independent basis of L over k . If we associate to each point P of V the point $Q = (f_0(P) : f_1(P) : \dots : f_n(P))$ of the n -dimensional projective space, then we obtain a rational mapping Φ_Σ from V to another variety V' . Outside the base points of Σ , the rational mapping Φ_Σ is regular; and the base points are the fundamental points of Φ_Σ . We say that Φ_Σ is the rational mapping defined by the linear system Σ . When Σ has no fixed components and Φ_Σ is a closed immersion, Σ is said to be **very ample** (or **ample**). For a divisor D , the set of positive divisors that are linearly equivalent to D is a linear system, which is called the **complete linear system** and is denoted by $|D|$. We usually write $l(D)$ instead of $l(|D|)$. If $|D|$ is very ample, we say that D is **very ample**. We say that D is **ample** (or **nondegenerate**) if mD is very ample for some $m > 0$.

0. Differential Forms

Let V be an n -dimensional irreducible variety, and let $A = k(V)$ be its function field. We denote by \mathfrak{D}^* the set of **derivations** of A over k , i.e., the k -linear mappings $D: A \rightarrow A$ satisfying $D(fg) = D(f)g + fD(g)$. Then \mathfrak{D}^* is an n -dimensional linear space over A . Let \mathfrak{D} denote the \dagger dual space of \mathfrak{D}^* over A . For each $f \in A$, let df be an element of \mathfrak{D} defined by $\langle df, D \rangle = D(f)$ ($D \in \mathfrak{D}^*$). Let x_1, x_2, \dots, x_n be a separating transcendence basis of A over k , in the sense that x_1, \dots, x_n are algebraically independent over k and A is a \dagger separable algebraic extension over $k(x_1, \dots, x_n)$. (Such a basis exists under the weaker hypothesis that k is \dagger perfect.)

Then dx_1, \dots, dx_n form a basis of \mathfrak{D} over A .

The homogeneous elements of degree r of the \dagger Grassmann algebra of \mathfrak{D} over A are called **differential forms of degree r** on V (or rational r -forms). The set of the differential forms of degree n is a 1-dimensional linear space over A spanned by $dx_1 \wedge \dots \wedge dx_n$.

A set of n functions f_1, \dots, f_n in A is called a system of **local coordinates** on an open set U of V if $f_1 - f_1(P), \dots, f_n - f_n(P)$ is a \dagger regular system of parameters of the local ring \mathfrak{O}_P for each $P \in U$. In that case, f_1, \dots, f_n is also a separating transcendence basis of A . If P is a simple point of V , then there exists a system of local coordinates on a suitable neighborhood of P . Let ω be a differential form of degree r on V , and write $\omega = \sum_{i_1 < \dots < i_r} \varphi_{(i)} df_{i_1} \wedge \dots \wedge df_{i_r}$, where (f_1, \dots, f_n) is a system of local coordinates around P . If the coefficients $\varphi_{(i)}$ are regular at P , then ω is said to be **regular** at P .

When V is a complete variety without singular points, a differential form that is everywhere regular on V is called a **differential form of the first kind** (or a **regular form**); the differential forms of the first kind are determined by the function field A and are independent of the choice of the nonsingular model V . The number of linearly independent differential forms of the first kind, of degree n , is denoted by p_g and is called the **geometric genus** of V .

Let V be a complete variety, W an irreducible subvariety of V of codimension 1, and P a point of W that is simple on V . Choose a system of local coordinates (f_i) . Given a differential form ω on V , we write it as a "polynomial" in the df_i , and denote by $v_W(\omega)$ the minimum of the values of the coefficients for the \dagger valuation $v_W(\cdot)$. The number $v_W(\omega)$ is determined by ω and W , and it is independent of the choice of P and of the local coordinates. Then ω defines a divisor $(\omega) = \sum_W v_W(\omega) W$ on V , which is called the **divisor of a differential form ω** . The divisor of a differential form of degree n ($= \dim V$) is called a **canonical divisor** and is usually denoted by K . The canonical divisors form a linear equivalence class of divisors.

P. Albanese Variety, Picard Variety, Néron-Severi Group

Let V be an irreducible variety. Then we can construct a couple (A, f) consisting of an \dagger Abelian variety called the **Albanese variety** of V and a rational mapping $f: V \rightarrow A$ (called an Albanese mapping) such that: (i) the image of f generates A , i.e., the sum of itself n times, $F: V^n \rightarrow A$, is generically surjective for sufficiently large n ; (ii) for every rational mapping $g: V \rightarrow B$ of V into an Abelian variety B ,

there exist a homomorphism $h: A \rightarrow B$ and a point $b \in B$ such that $g = h \cdot f + b$. The Albanese variety is uniquely determined up to isomorphisms and f is determined up to translations.

In the case of $k = \mathbb{C}$, if V is a complete nonsingular variety and if q is the dimension of the linear space of differential 1-forms of the first kind on V , then the first \dagger Betti number $i_1^?$, is equal to $2q$. Let $\omega_1, \dots, \omega_q$ be a basis of the linear space and let $\gamma_1, \dots, \gamma_{2q}$ be a basis of the first homology group modulo torsion. Put $\alpha_{ji} = \int_{\gamma_j} \omega_i$ and $\alpha_j = (\alpha_{j1}, \dots, \alpha_{jq})$. Then the period vectors α_j ($1 \leq j \leq 2q$) are linearly independent over \mathbb{R} in \mathbb{C}^q . If Γ denotes the discrete subgroup of \mathbb{C}^q generated by the α_j , then the quotient group \mathbb{C}^q/Γ is the Albanese variety of V . The Albanese mapping is given by the mapping $P \rightarrow (\int_Q^P \omega_1, \dots, \int_Q^P \omega_q) \pmod{\Gamma}$, where P is a variable point on V and Q is a fixed point on V (\rightarrow 232 Kähler Manifolds C).

Replacing the term "rational mapping" by "morphism" in the definition, we can define the **strict Albanese variety** of V and prove its existence. It is a quotient Abelian variety of the Albanese variety of V . If V is nonsingular, both coincide by virtue of the \dagger strong minimality of an Abelian variety (\rightarrow Section I).

Let V be a complete normal variety, U the set of the simple points of V , and D a divisor on V . Then D is said to be **algebraically equivalent to 0** if there exist a nonsingular curve C , a divisor Γ on $U \times C$, and two points P and Q on C such that D can be written as $D = \varphi_P^*(\Gamma) - \varphi_Q^*(\Gamma)$, where φ_P and φ_Q are the morphisms $\varphi_P: U \rightarrow U \times P \rightarrow U \times C$ and $\varphi_Q: U \rightarrow U \times Q \rightarrow U \times C$. We denote by $\mathfrak{G}(V)$, $\mathfrak{G}_a(V)$, and $\mathfrak{G}_l(V)$ the set of all divisors on V , the set of divisors that are algebraically equivalent to 0, and the set of divisors that are linearly equivalent to 0, respectively. We can introduce a canonical structure of an Abelian variety into $\mathfrak{G}_a(V)/\mathfrak{G}_l(V)$, which is called the **Picard variety** of V . The dimension q of the Picard variety is called the **number of irregularity** of V ; if $q = 0$ we say that V is **regular**.

The Albanese variety and the Picard variety of V are \dagger isogeneous to each other, and each one is the Picard variety of the other. If V is a curve, they are isomorphic and called the **Jacobian variety** (\rightarrow 9 Algebraic Curves E).

Using Cartier divisors instead of divisors, we get an analogous theory to construct another kind of Picard variety which turns out to be isomorphic to the Picard variety of the strict Albanese variety of V . The group of the linear equivalence classes of Cartier divisors can be identified with $H^1(V, \mathcal{O}_V^*)$, where \mathcal{O}_V^* is the sheaf of multiplicative groups of the invertible elements in \mathcal{O}_V . From this point of view, we can generalize the theory of Picard variety to the case of schemes also. The theory thus obtained

is called the theory of **Picard schemes** [60].

The quotient group $\text{NS}(V) = \mathfrak{G}(V)/\mathfrak{G}_a(V)$ is finitely generated [66] and is called the **Neron-Severi group** of V . We call the rank of $\text{NS}(V)$ the **Picard number** of V and denote it by $p(V)$. In the case of a nonsingular projective variety over $k = \mathbb{C}$ we have an inequality $p(V) \leq h^{1,1}(V) (= \dim, H^1(V, \Omega_V^1))$ and the **Lefschetz number** $B_2(V) - \rho(V)$ is a birational invariant (where $B_2(V)$ is the second Betti number of V) [39]. For the positive characteristic case, however, the above inequality does not hold in general [61].

The torsion part of $\text{NS}(V)$ is $\mathfrak{G}_n(V)/\mathfrak{G}_a(V)$, where $\mathfrak{G}_n(V)$ denotes the group of divisors numerically equivalent to zero (\rightarrow Section Q) (T. Matsusaka). The last fact cannot be generalized for higher codimensional cycles [25].

Q. General Intersection Theory

Let V be an irreducible variety of dimension n , and let A and B be irreducible subvarieties of V of dimension r and s , respectively. If C is a proper component of $A \cap B$, we can define the **intersection multiplicity** $i(A, B, C; V)$ of A and B along C on V , which has properties consistent with our geometric intuitions. In particular, it is invariant under biregular mappings. If A and B intersect properly on V and if C_1, \dots, C_N are the proper components of $A \cap B$, we define an $(r+s-n)$ -cycle $A \cdot B$ by $A \cdot B = \sum_v i(A, B, C_v; V) C_v$ and call it the **intersection product** of A and B . If each component X_α of an r -cycle $X = \sum_\alpha n_\alpha X_\alpha$ and each component Y_β of an s -cycle $Y = \sum_\beta m_\beta Y_\beta$ intersect properly, we define

$$X \cdot Y = \sum_\alpha \sum_\beta n_\alpha m_\beta X_\alpha \cdot Y_\beta.$$

Then we have the associative law $(X \cdot Y) \cdot Z = X \cdot (Y \cdot Z)$ for cycles X , Y , and Z whenever both sides are defined. Two r -cycles X_1 and X_2 are said to be **numerically equivalent** if for every $(n-r)$ -cycle Y that intersects them properly we have $\deg(X, Y) = \deg(X_2, Y)$.

The theory of intersection is one of the most basis theories in algebraic geometry, for the other theories can be constructed from it [92].

R. Chow Rings

Let U and T be nonsingular irreducible varieties. If Z is a cycle on $U \times T$ such that $Z \cdot (U \times t)$ is defined for every point t of T , we put $Z \cdot (U \times t) = X(t) \times t$, and we obtain a family $\{X(t)\}$ of cycles on U parametrized by the points of T . Such a family is called an **algebraic family** of cycles. Two cycles X_1 and X_2 whose difference is equal to the difference of two cycles in an algebraic family are said to

be **algebraically equivalent**. In particular, if $X_1 - X_2$ can be expressed as the difference of two cycles of an algebraic family parametrized by the points of the affine line, then X_1 and X_2 are said to be **rationally equivalent**. The set of the cycles that are algebraically (rationally) equivalent to 0 is a subgroup $\mathfrak{Z}_a(U)$ ($\mathfrak{Z}_{rat}(U)$) of $J(U)$. For divisors, rational equivalence coincides with linear equivalence.

Let $f: V \rightarrow U$ be a morphism between nonsingular irreducible varieties. For an irreducible subvariety W of V , let W' denote the closure of $f(W)$, and put $f_3(W) = 0$ if $\dim W > \dim W'$ and $f_3(W) = mW'$ if $\dim W = \dim W'$, where $m = [k(W): k(W')]$ is the degree of the morphism $W \rightarrow W'$. Extending f_3 by linearity, we obtain a module homomorphism $f_3: \mathfrak{Z}(V) \rightarrow \mathfrak{Z}(U)$. If f is proper, then f_3 induces a module homomorphism f_* from $\mathfrak{Z}(V)/\mathfrak{Z}_{rat}(V) = A(V)$ to $A(U)$.

Let U and V be nonsingular irreducible projective varieties, and $f: V \rightarrow U$ be a morphism with graph Γ . If $Y \in \mathfrak{Z}(U)$ is such that $\Gamma \cdot (V \times Y)$ is defined, we denote by $f_3^*(Y)$ the image of $\Gamma \cdot (V \times Y)$ under the induced isomorphism $\Gamma \rightarrow V$. Each class of the rational equivalence class group $A(U) = \mathfrak{Z}(U)/\mathfrak{Z}_{rat}(U)$ contains a cycle for which f_3^* is defined. Hence we can define $f^*: A(U) \rightarrow A(V)$. Let $\Delta: U \rightarrow U \times U$ be the diagonal morphism, and define $x \cdot y = \Delta^*(x \times y)$ for $x, y \in A(U)$. Then $A(U)$ is a ring with respect to this product; moreover, it is a \dagger graded ring with the grading by codimension. This graded ring $A(U)$ is called the **Chow ring** of U , and the mapping $f^*: A(U) \rightarrow A(V)$ is a ring homomorphism. If f is proper, we have $f_*(y \cdot f^*(x)) = f_*(y) \cdot x$ for $x \in A(U)$, $y \in A(V)$.

Zero Cycles. Let V be a nonsingular irreducible projective variety over an uncountable algebraically closed field of characteristic zero (say C). Denote by $A_0(V)$ the group of classes of 0-cycles of degree 0 on V modulo rational equivalence. We say $A_0(V)$ is finite-dimensional if the mapping $V^n \times V^n \rightarrow A_0(V)$ sending $(a_1, \dots, a_n, b_1, \dots, b_n)$ to $\sum a_i - \sum b_i$ is surjective for a certain n .

If $A_0(V)$ is finite-dimensional, then the canonical mapping $A_0(V) \rightarrow \text{Alb}(V)$ induced from $V \rightarrow \text{Alb}(V)$ is bijective (A. A. Rojzman). In general, the torsion part of $A_0(V)$ is isomorphic to that of $\text{Alb}(V)$ (Rojzman [75]). If $h^{p,0}(V) > 0$ for some $p > 1$, then $A_0(V)$ is not finite-dimensional (Mumford, Rojzman [74]). If V is a surface which is not of general type and with $h^{2,0}(V) = 0$, then $A_0(V)$ is isomorphic to $\text{Alb}(V)$ (S. Bloch, A. S. Kas, D. Lieberman [13]). There exist surfaces of general type with $h^{2,0} = 0$ such that $A_0(V) = 0$. For instance, Godeaux surfaces are such surfaces (H. Inose [43]).

One Cycles. Let V be a nonsingular projective irreducible variety and $N(V)$ denote the numerical equivalence classes of 1-cycles on V ($\otimes_{\mathbb{Z}} \mathbb{R}$). Via the intersection pairing, $N(V)$ is dual to the Neron-Severi group tensored by \mathbb{R} (Section P); and let $NE(V)$ denote the smallest cone in $N(V)$ containing all effective 1-cycles. S. Mori studied the structure of $NE(V)$ in detail. One of his results is stated as follows: If the anticanonical divisor is ample, then there exist rational irreducible curves l_1, \dots, l_r such that (1) $-(l_i \cdot K_V) < n + 1$ and (2) $NE(V) = \mathbb{R}_+[l_1] + \dots + \mathbb{R}_+[l_r]$, where $n = \dim V$, K_V is the canonical divisor on V , and $[l_i]$ denotes the class represented by l_i . Moreover, Mori has proved that (1) if K_V is not numerically effective, e.g., if $-K_V$ is ample, then V contains a rational curve, and (2) if the tangent bundle of V is \dagger ample, then V is isomorphic to the projective space [59]. (This is called the **Hartshorne conjecture**.)

S. Chow Coordinates, Hilbert Schemes

Consider an irreducible algebraic correspondence T between irreducible varieties V and W . Let V' and W' be the closed images of the projections of T to V and W , and let a and c be their dimensions, respectively. Take generic points P, Q of V', W' , and consider the total transforms $T\{P\}$ of P to W and $T\{Q\}$ of Q to V . Denoting the dimensions of $T\{P\}$ and $T\{Q\}$ by b and d , respectively, we have $a + b = c + d$, where both sides are equal to the dimension of T . This property is called the **principle of counting constants**.

This simple principle has wide application. For instance, let V be an r -dimensional variety in $\mathbb{P}^n(k)$, and let $\sum_j u_{ij} X_j = 0$ ($0 \leq i \leq r$) be the equations of $r + 1$ hyperplanes H_i . The condition $V \cap (H_0 \cap \dots \cap H_r) \neq \emptyset$ defines an irreducible algebraic correspondence T between V and the multiprojective space $W = F^{r+1}(k) \times \mathbb{P}^n(k)$ with u_{ij} as coordinates: $T = \{(x, u) \mid x \in V, \sum u_{ij} X_j = 0\}$. In this case, we have $a = r, b = (n - 1)(r + 1)$, and $d = 0$ in the notation introduced previously, so that $c = n(r + 1) - 1$. This implies that W' is of codimension 1 in W ; hence W' is defined by a single equation $F(u_{ij}) = 0$. This form F is the **associated form** of V of B. L. van der Waerden and W. L. Chow. It is a homogeneous form of degree d ($d = \deg(V)$) in each (u_{i0}, \dots, u_{in}) and is symmetric in the indices i . More generally, for a positive cycle $X = \sum n_x V_x$ of dimension r and of degree d ($= \sum n_x \deg(V_x)$) in $\mathbb{P}^n(k)$, the product $\prod F_x^{n_x}$ of the associated forms F_x of V_x is called the **associated form** of X . The coefficients of F , arranged in a fixed order and regarded as the homogeneous coordinates of a point of a pro-

jective space, are called the **Chow coordinates** of X . This is a natural generalization of the \dagger Plücker coordinates. Given r , d and a projective variety $U (\subset \mathbf{P}^n(k))$, the set of Chow coordinates of the positive cycles that are contained in U , whose dimension is r and whose degree is d , is a projective variety called a **Chow variety**.

In scheme theory, the Hilbert scheme is introduced in the following way [60]. Let \mathbf{P}_Z^n be the n -dimensional projective space over $\text{Spec } \mathbf{Z}$. For a locally Noetherian scheme S , and for a closed subscheme Z of $\mathbf{P}_Z^n \times S$, we have the composition of the immersion of Z and the projection $\mathbf{P}_Z^n \times S \rightarrow S$; this is denoted by $f_Z: Z \rightarrow S$. For any point s of S , $f_Z^{-1}(s)$ is a closed subscheme of $\mathbf{P}_{k(s)}^n$ and the restriction of $\mathcal{O}(1)$ to $f_Z^{-1}(s)$ is denoted by L_s . Put $\mathcal{M}(S) = \{Z \subset \mathbf{P}_Z^n \times S \mid f_Z \text{ is flat}\}$. Then M becomes a contravariant functor to the category of sets, which is representable, i.e., there exists a locally Noetherian scheme M such that \mathcal{M} is naturally isomorphic to the functor $\text{Hom}(-, M)$. M is written as $\text{Hilb}(\mathbf{P}_Z^n)$ and is called the **Hilbert scheme**. There exists a closed subscheme W of $\mathbf{P}_Z^n \times \text{Hilb}(\mathbf{P}_Z^n)$ with flat f_W such that for any $Z \in M(S)$, there exists a unique $\varphi: S \rightarrow M = \text{Hilb}(\mathbf{P}_Z^n)$ in such a way that $Z = W \times_M S$. In particular, if X is a closed subscheme of \mathbf{P}_Z^n over a field k , there exist $\text{SE } \text{Hilb}(\mathbf{P}_Z^n)$ and a field extension $k/k(s)$ such that $X = f_W^{-1}(s) \otimes_{k(s)} k$. Thus $\text{Hilb}(\mathbf{P}_Z^n)$ parametrizes all closed subschemes of \mathbf{P}_k^n . Fix a polynomial P and define $\mathcal{M}^P(S)$ to be $\{Z \in M(S) \mid \chi(L_s^{\otimes m}) = P(m) \text{ for all } s\}$. \mathcal{M}^P is also a contravariant functor, represented by a scheme $\text{Hilb}^P(\mathbf{P}_Z^n)$ which is projective over $\text{Spec } \mathbf{Z}$. The direct sum of all $\text{Hilb}^P(\mathbf{P}_Z^n)$ is just $\text{Hilb}(\mathbf{P}_Z^n)$.

T. Algebraic Geometry and Complex Analytic Geometry

When $k = \mathbf{C}$, an algebraic variety is called a **complex algebraic variety**, and it has the structure of a \dagger complex analytic manifold or (if it has singular points) of a \dagger analytic space. If we denote by $\mathcal{O}_{V,x}^h$ the ring of holomorphic functions at a point x of V , then $\mathcal{O}_{V,x} \subset \mathcal{O}_{V,x}^h$, and their \dagger completions coincide. If x is a simple point of V , then $\mathcal{O}_{V,x}^h$ is the \dagger ring of convergent power series, and its completion is the \dagger ring of formal power series. The prime ideals of $\mathcal{O}_{V,x}^h$ remain prime under completion (M. Nagata). Because of this, the analytic behavior of V in a neighborhood of x can be investigated algebraically through the completion of $\mathcal{O}_{V,x}$. If V is complete, then the analytic coherent sheaves on V and the algebraic coherent sheaves on V correspond to each other bijectively; consequently, for propositions that can

be stated in terms of coherent sheaves, the results in the analytic sense remain valid in the algebraic sense also, and vice versa (J.-P. Serre [82]) (- 72 Complex Manifolds E).

U. Topology of Algebraic Varieties

Every algebraic variety defined over \mathbf{R} (or \mathbf{C}) can be triangulated by real analytic cells [46]. Let V be a nonsingular connected algebraic variety defined over \mathbf{C} . For an algebraic automorphism σ of \mathbf{C} we can define V^σ by letting σ operate on the coefficients of the defining equations of open affine coverings of V . V and V^σ are not necessarily homeomorphic or even of the same homotopy type [83]. Grothendieck has shown that there is a \dagger spectral sequence $E_1^{p,q} = H^p(V, \Omega_V^q) \Rightarrow H^{p+q}(V, \mathbf{C})$, called a **Hodge spectral sequence**. This spectral sequence can be defined in a purely algebraic way and $H^{p+q}(V, \mathbf{C})$ can be considered as the \dagger hypercohomology of the de Rham complex $\{\Gamma(V, \Omega_V^q), d\}$ of V . If V is projective, V carries a Kähler metric and by the theory of harmonic integrals the Hodge spectral sequence degenerates and $E_1^{p,q}$ is \mathbf{C} -isomorphic to the complex conjugate of $E_1^{q,p}$ (- 232 Kähler Manifolds B). This is also the case if V is complete.

The topology of a nonsingular projective surface was studied by Lefschetz using the method of Lefschetz pencils. For a projective nonsingular irreducible variety V of dimension n defined over \mathbf{C} , a **Lefschetz pencil** $\{W_t\}_{t \in \mathbf{P}^1}$ of V is, by definition, a linear pencil consisting of hyperplane sections W_t of V such that: (i) for all $t \in U = \mathbf{P}^1 - \{t_1, t_2, \dots, t_d\}$, W_t is nonsingular; (ii) each W_{t_i} has only one singular point that is an ordinary double point; and (iii) $W_0 \cap W_\infty$ is nonsingular, where we assume $0, \infty \in U$. Embed V into \mathbf{P}^N by a high multiple of a hyperplane of V and take a generic linear pencil $\{H_t\}$ of hyperplanes in \mathbf{P}^N . Then $\{W_t = H_t \cap V\}$ is a Lefschetz pencil of V . By blowing up V along $W_0 \cap W_\infty$ we obtain a smooth variety \tilde{V} and a surjective morphism $\pi: \tilde{V} \rightarrow \mathbf{P}^1$. Let $W = \pi^{-1}(0)$, $\pi_U = \pi|_{\pi^{-1}(U)}$. $R^p \pi_{U*} \mathbf{Q}$ is a local system attached to the monodromy representation $\varphi_p: \pi_1(U, 0) \rightarrow GL(H^p(W, \mathbf{Q}))$. φ_p is trivial if $p \neq n-1$. For each point t_i there corresponds a cocycle δ_i of $H^{n-1}(W, \mathbf{Q})$ called a **vanishing cocycle** such that if γ_i is a loop based at 0 going once (counterclockwise) around t_i , we have, for each $x \in H^{n-1}(W, \mathbf{Q})$, $\varphi_p(\gamma_i)(x) = x \pm (x, \delta_i) \delta_i$, where $(\ , \)$ is the intersection pairing of $H^{n-1}(W, \mathbf{Q})$. $\varphi_p(\gamma_i)$ is called a **Picard-Lefschetz transformation**. The main results due to Lefschetz are restated as follows. (1) (**Weak Lefschetz theorem**). The natural homomorphism $H^i(V, \mathbf{Q}) \rightarrow H^i(W, \mathbf{Q})$ is an isomorphism for $0 \leq i \leq n-2$ and is an injection for

$i = n - 1$, or equivalently $H_i(V, \mathbf{W}, \mathbf{Q}) = 0$ for $0 \leq i \leq n - 1$. (2) (Strong Lefschetz theorem). Let ξ be the cohomology class of $H^2(V, \mathbf{Q})$ corresponding to the hyperplane section \mathbf{W} , and let $L: H^*(V, \mathbf{Q}) \rightarrow H^{*+2}(V, \mathbf{Q})$ be the homomorphism defined by the cup product with ξ . Then for each $i \leq n$, $L^{n-1}: H^i(V, \mathbf{Q}) \rightarrow H^{2n-i}(V, \mathbf{Q})$ is an isomorphism. The weak Lefschetz theorem is true for a cohomology with integral coefficients. In fact, V-W has the homotopy type of a real n -dimensional finite CW complex, and $\pi_r(V, \mathbf{W}) = 0$ for $r < n$. The strong Lefschetz theorem is equivalent to the statement that $H^{n-1}(W, \mathbf{Q})$ is the direct sum of the vector space spanned by the vanishing cocycles δ_i and the vector space spanned by the invariant cocycles (i.e., $\varphi_{n-1}(\gamma_i)x = x, i = 1, \dots, d$). Lefschetz's original proof of this statement is incomplete, and no direct topological proof is known. The transcendental proof of (2) is given by the theory of harmonic integrals. A version of Lefschetz pencils is a proper morphism $f: X \rightarrow D = \{z \mid |z| < \varepsilon\}$ of a complex manifold X onto a disk D such that $f^* = f \circ f^{-1}(D^*)$, $D^* = D - \{0\}$, is of maximal rank at every point of $f^{-1}(D^*)$. Fix a point $s \in D^*$. $\pi_1(D^*, s)$ operates on $H^j(W, \mathbf{Z}, \mathbf{W} = f^{-1}(s))$, and we have a representation $\varphi_j: \pi_1(D^*, s) \rightarrow GL(H^j(W, \mathbf{Z}))$. For a loop γ based at s and going once around 0 the Picard-Lefschetz transformation $\varphi_j(\gamma)$ is quasiunipotent (i.e., for a certain integer m , $\varphi_j(\gamma)^m$ is unipotent).

For a nonsingular projective variety defined over a field k with characteristic $p > 0$, the above Lefschetz theorems hold for an l -adic cohomology ($l \neq p$) [19, 30 (SGA 7)] (Section AA). Using the theory of finite étale coverings of an algebraic variety defined over a field k , we can define the algebraic fundamental group and the algebraic homotopy groups, which are profinite completions of the topological fundamental group and the topological homotopy groups, respectively, where $k = \mathbb{C}$ [5, 30 (SGA 1)]. Let (X, x) be a germ of a complex space with isolated singular point x . (X, x) is always algebraizable, i.e., the completion $\hat{\mathcal{O}}_{X, x}$ of the analytic local ring $\mathcal{O}_{X, x}$ is isomorphic to the completion of the local ring of a closed point of an algebraic variety defined over \mathbb{C} [3]. For topological type of algebraic surfaces \rightarrow Moishezon [57].

V. Hodge Theory

Let $H_{\mathbf{R}}$ be a finite-dimensional real vector space containing a lattice $H_{\mathbf{Z}}$, and let $\mathbf{H} = H_{\mathbf{R}} \otimes_{\mathbf{R}} \mathbf{C}$ be its complexification. A Hodge structure of weight m on \mathbf{H} (or $H_{\mathbf{R}}$) is, by definition, a direct sum decomposition $\mathbf{H} = \bigoplus_{p+q=m} H^{p,q}$, $\bar{H}^{p,q} \cong H^{q,p}$, where $H^{p,q}$ is a

complex vector subspace and the overbar denotes complex conjugation. If \mathbf{H} and \mathbf{H}' carry Hodge structures of weight m and m' , respectively, then $\mathbf{H} \otimes \mathbf{H}'$, $\text{Hom}_{\mathbf{C}}(\mathbf{H}, \mathbf{H}')$, $\wedge^p \mathbf{H}$, and \mathbf{H}^* carry Hodge structures of weight $m + m'$, $m' - m$, pm , and $-m$, respectively. For a Hodge structure \mathbf{H} of weight m , $F^p \mathbf{H} = \bigoplus_{k \geq p} H^{k, m-k}$, $p = 0, 1, \dots, m$, induces a decreasing filtration. Let \mathbf{H} be a Hodge structure of weight m , and let Q be a bilinear form on \mathbf{H} . If the following three conditions are satisfied, the Hodge structure \mathbf{H} is said to be polarized by Q . (i) Q is defined over \mathbf{Q} and is symmetric (skew-symmetric) if m is even (odd). (ii) $Q(H^{p,q}, H^{p',q'}) = 0$ unless $p = p', q = q'$. (iii) $(\sqrt{-1})^{p-q} Q(v, \bar{v}) > 0$ for nonzero $v \in H^{p,q}$. Let \mathbf{V} be a compact Kähler manifold. Then $\mathbf{H} = H^*(V, \mathbb{C})$ carries the Hodge structure induced by the type (p, q) -decomposition (\rightarrow 232 Kähler manifolds B). This is also the case if \mathbf{V} is a compact complex manifold which is the image of a holomorphic mapping from a compact Kähler manifold of the same dimension. Moreover, if \mathbf{V} is projective, the Hodge-Riemann bilinear relations define a natural polarization on the subspace P of \mathbf{H} consisting of all primitive cohomology classes.

Each algebraic cycle \mathbf{W} of codimension s determines a cohomology class $[\mathbf{W}] \in H^{2s}(V, \mathbf{Q})$, which belongs to $H^{s,s}(V)$. (Such a class is called an algebraic cycle.) The converse of this fact is called the Hodge conjecture, which says that $H^{2s}(V, \mathbf{Q}) \cap H^{s,s}(V)$ is spanned by algebraic cycles. The case $s = 1$ has been verified by Hodge, Lefschetz, and Kodaira.

Let $\mathbf{V}(\mathbf{W})$ be a smooth irreducible algebraic variety defined over \mathbb{C} (complex manifold), and let $\varphi: V \rightarrow \mathbf{W}$ be a projective smooth morphism with connected fibers. Then $\mathbf{H} = R^m \varphi_* \mathbf{C}$ is a flat vector bundle on \mathbf{W} with the flat connection ∇ . \mathbf{V} is often called the Gauss-Manin connection, and it can be defined algebraically if \mathbf{W} is also algebraic.

For each fiber $V_s = \varphi^{-1}(s)$, $s \in \mathbf{W}$, the filtration $F^p H^m(V_s, \mathbb{C}) = \bigoplus_{k \geq p} H^{k, m-k}(V_s)$ induces a complex subbundle \mathbf{F}^p and the connection ∇ has the property $\nabla(\mathcal{O}(\mathbf{F}^p)) \subset \mathcal{O}(\mathbf{F}^{p-1} \otimes \mathbf{T}^*)$, where \mathbf{T} is the tangent bundle of \mathbf{W} [26]. Moreover, if \mathbf{W} is algebraic, \mathbf{V} is a differential equation with regular singular points on \mathbf{W} where \mathbf{W} is a smooth compactification of \mathbf{W} , such that $\mathbf{W} - \mathbf{W}$ is a divisor with normal crossings (\rightarrow Section L). If we consider the subbundle \mathbf{P} of \mathbf{H} consisting of all primitive cohomology classes, the polarization on each fiber induces a Hermitian pseudometric on \mathbf{P} . Curvatures of bundles $\mathbf{P} \cap \mathbf{F}^p$ have been studied by Griffiths [26]. There exists a classifying space D for polarized Hodge structures and there exists a holomorphic mapping of the universal covering \tilde{W} of \mathbf{W} into D , usually

called a period mapping. D may not be a bounded symmetric domain but has several interesting properties [24, 27]. In some cases D/Γ with a suitable discrete subgroup Γ is the moduli space of polarized algebraic varieties (e.g., curves, Abelian varieties).

P. Deligne [18] has generalized Hodge theory to arbitrary algebraic varieties (more generally schemes of finite type over \mathbb{C}). The simplest case is the Hodge theory of a smooth noncomplete irreducible variety X . By Nagata's embedding theorem [63] there exists a complete algebraic variety \bar{X} such that $Y = \bar{X} - X$ is a subvariety. By virtue of Hironaka's resolution theorem we can assume that \bar{X} is nonsingular and that Y is a divisor with normal crossings. Let $\Omega_X^1(\log Y)$ be a sheaf of germs of meromorphic 1-forms with logarithmic pole along Y , i.e., locally written as $\sum_{i=1}^k a_i(x)(dx_i/x_i) + \sum_{j=k+1}^n a_j(x)dx_j$, where (x_1, \dots, x_n) is a system of local coordinates with center $p \in Y$ in X such that $x_k = 0$ is a local equation of Y and $a_i(x), a_j(x)$ are holomorphic at p . Using the complex $\{\Omega_X^p(\log Y) = \bigwedge^p \Omega_X^1(\log Y), d\}$, with a suitable filtration, Deligne has shown that $H = H^m(X, \mathbb{C})$ carries a mixed Hodge structure and this structure is independent of the choice of X . The mixed Hodge structure on H consists of two finite filtrations, i.e., $\text{Oc.} \subset W_{n-1} \subset W_n \subset \dots \subset H$, the weight filtration which is defined over \mathbb{Q} , and $\dots \subset F^p \subset F^{p+1} \subset \dots \subset H$, the Hodge filtration such that F^p induces on W_n/W_{n-1} a Hodge structure of weight n . As a corollary he has shown that a meromorphic p -form on \bar{X} with logarithmic pole along Y (i.e., a section of $\Omega_X^p(\log Y)$) is d -closed on X , and $\omega = 0$ if and only if $\omega|_X$ is zero in $H^p(X, \mathbb{C})$. An important application of the theory of this mixed Hodge structure on $H^m(X, \mathbb{C})$ is the following. Let V and W be smooth irreducible varieties, and let $\varphi: V \rightarrow W$ be a smooth projective morphism. If \bar{V} is a smooth compactification of V , the canonical homomorphism $H^m(V, \mathbb{Q}) \rightarrow H^0(W, R^m\varphi_*\mathbb{Q})$ is surjective. Fix a point $s \in W$. $\pi_1(W, s)$ operates on $H^m(V_s, \mathbb{Q})$. Then this action is semisimple [18].

W. Deformations, Moduli, Algebraic Spaces

In this section for simplicity the field k is assumed to be algebraically closed. Let X be an algebraic scheme over k . A (flat) **deformation of X over a connected scheme S** over k with base point s_0 consists of the following data: (1) A morphism $p: \mathfrak{X} \rightarrow S$ that is flat and of finite type. If X is complete, p is also proper. (2) A closed point $s_0 \in S$ such that the fiber $\mathfrak{X} \times_S k(s_0)$ is isomorphic to X . For any closed point $s \in S$, the fiber $X_s = \mathfrak{X} \times_S k(s)$ is called a **flat deforma-**

tion of X . If X is smooth and complete, we assume further that p is smooth. Similarly, we can define a deformation of a polarized algebraic manifold, an embedding deformation of X in an algebraic scheme Y over k , a deformation of an affine scheme with isolated singular points, and a deformation of vector bundles on a fixed algebraic scheme over k , etc. The theory has two aspects: local theory and global theory.

Let (R, \mathfrak{m}) be a complete Noetherian local ring such that $R/\mathfrak{m} = k$. Set $R_n = R/\mathfrak{m}^n$. A formal deformation X_R of X is a sequence $\{X_n\}$ such that (i) X_n is a deformation of X over $\text{Spec}(R_n)$ and (ii) there is a compatible sequence of isomorphisms $X_n \otimes_{R_n} R_{n-1} \xrightarrow{\sim} X_{n-1}$ for any n . Let (FLA/k) be the category of finite-dimensional commutative local k -algebras. The local theory of deformation is the study of the covariant functor F of (FLA/k) to (Set) , where, for $A \in (\text{FLA}/k)$, $F(A)$ is the set of isomorphism classes of deformations of X over $\text{Spec}(A)$. The functor is in general neither representable nor prorepresentable (i.e., there exists a formal deformation X_R of X such that $F(A) = \text{Hom}_{k\text{-alg}}(R, A)$). But, under reasonably mild conditions on F , F has the hull R [78]. That is, there is a formal deformation X_R of X and a natural transformation $j: G \rightarrow F$, where $G(A) = \text{Hom}_{k\text{-alg}}(R, A)$, such that j is formally smooth (i.e., for any surjection $A' \twoheadrightarrow A$ in (FLA/k) , $G(A') \rightarrow G(A) \times_{F(A)} F(A')$ is surjective) and $G(k[\varepsilon]) \rightarrow F(k[\varepsilon])$ is bijective for the ring of dual numbers $k[\varepsilon]$. The formal deformation X_R is called a **versal deformation** of X . The hull R is unique up to noncanonical isomorphism. The deformation functor F has the hull R if X is (i) a complete algebraic scheme over k , (ii) an affine scheme with isolated singular points, (iii) a polarized algebraic variety over k , or (iv) a vector bundle on a complete algebraic scheme over k , etc. If there exists a deformation $\pi: \mathfrak{X} \rightarrow S$ of X over a scheme S with base point s_0 over k such that $R = \hat{\mathcal{O}}_{S, s_0}$, $X_n \cong \mathfrak{X} \otimes_S R_n$, the formal deformation $\{\pi_n: X_n \rightarrow \text{Spec}(R_n)\}$ is called **algebraizable**. Algebraizability of the versal deformation has been studied by M. Artin. Since the assumption that S is a scheme is rather restrictive, Artin has introduced the notion of algebraic spaces and has considered algebraizability in the category of algebraic spaces [2–4]. For a complete algebraic variety, the versal deformation is not necessarily algebraizable and we need to consider deformations of polarized algebraic varieties. The versal deformation of an affine variety with an isolated singularity is algebraizable in the category of algebraic spaces. For the global theory of deformations, we need the projectivity assumption, and the theory is essentially reduced to the theory of

Hilbert schemes (or Chow varieties) (\rightarrow Section S). The problem of moduli is considered as the study of the set A_4 of all isomorphism classes of deformations of X . Usually we consider the moduli of polarized varieties.

Let (Sch) be the category of Noetherian schemes and \mathcal{M} a contravariant functor from (Sch) to (Set) defined by $\mathcal{M}(X) = \{\text{isomorphism classes of families of polarized varieties parametrized by } X \in \text{Ob}(\text{Sch}), \text{ e.g., families of polarized Abelian varieties with(out) additional conditions}\}$. The functor \mathcal{M} is called a moduli functor. If it is represented by a scheme M , \mathcal{M} is called a **fine moduli scheme** [60]. In this case there is the universal family $\pi: \mathcal{V} \rightarrow M$. In many cases the moduli functor is not representable. A **coarse moduli scheme** for a given moduli problem is a scheme A_4 together with a natural transformation $\phi: \mathcal{M} \rightarrow \text{Hom}(-, M)$ such that (1) $\phi(\text{Spec}(k)): M(\text{Spec}(k)) \rightarrow \text{Hom}(\text{Spec}(k), M)$ is bijective for any algebraically closed field k , and (2) for any scheme N and any natural transformation $\psi: \mathcal{M} \rightarrow \text{Hom}(-, N)$, there is a unique natural transformation $\lambda: \text{Hom}(-, M) \rightarrow \text{Hom}(-, N)$ with $\psi = \lambda \cdot \phi$. A coarse moduli scheme is called a **moduli space** or a **moduli scheme**.

In many cases, the moduli space can be obtained as the quotient space of a certain (locally closed) subset H of a Hilbert scheme by the following equivalence relation: $s \sim s'$ $\Leftrightarrow H$ if and only if $X_s \cong X_{s'}$ as polarized varieties, where $\pi: \tilde{X} \rightarrow H$, $\pi^{-1}(s) = X_s$ (T. Matsusaka [53]). This equivalence relation is often induced by an action of a \dagger reductive algebraic group G . Suppose that a reductive algebraic group G operates on an algebraic k -scheme Z . A G -invariant morphism $f: Z \rightarrow Y$ (i.e., for the trivial action of G on Y , f is a G -equivariant morphism) is called a **geometric quotient** if (1) f is a surjective affine morphism and $f_*(\mathcal{O}_Z)^G = \mathcal{O}_Y$, (2) if X is a G -stable closed subset of Z , then $f(X)$ is closed in Y , and (3) for $x_1, x_2 \in Z$, $f(x_1) = f(x_2)$ if and only if the G -orbits of x_1 and x_2 are the same. Let G be a reductive algebraic group, $\chi: G \rightarrow \text{Aut}(V)$ a rational representation on a finite-dimensional vector space V over k , and $v_0 \neq 0$ a G -invariant point. Then there exists a G -invariant homogeneous polynomial F of degree ≥ 1 on V such that $F(v_0) \neq 0$ (W. J. Haboush [32]). This implies that if a reductive group G operates on an algebraic k -affine scheme $\text{Spec}(A)$, then the invariant ring A^G is a finitely generated k -algebra and, moreover, if any G -orbit in $\text{Spec}(A)$ is closed, the natural morphism $\text{Spec}(A) \rightarrow \text{Spec}(A^G)$ is a geometric quotient. For a quasiprojective scheme Z over k with an action of a reductive group G we need a notion of stable points [60, 62]. The subset of stable points of Z consists of all geometric

points of a G -stable open subscheme Z^s of Z , and there exists a geometric quotient $f: Z^s \rightarrow Y$ where Y is quasiprojective (Mumford [60], C. S. Seshadri [79], Haboush [32]). In this way Mumford has shown the existence of coarse moduli schemes of nonsingular complete irreducible algebraic curves and polarized Abelian varieties. But, in general, analysis of stable points is very difficult and it is desirable to extend the category of schemes so that it becomes easier to obtain a quotient. Matsusaka has introduced the notion of a Q -variety [52]. M. Artin has introduced the notion of an algebraic space (D. Knutson [49]).

An **algebraic space** X of **finite type** consists of an affine scheme U and a closed subscheme $R \subset U \times U$ such that (1) R is an \dagger equivalence relation, and (2) the projections $p_i: R \rightarrow U$ ($i = 1, 2$) are étale. (These are often written as $R \rightrightarrows U \rightarrow X$.) A morphism $g: V \rightarrow X$ of an affine scheme V to an algebraic space X consists of a closed subscheme $W \subset U \times V$ such that (1) the projection $W \rightarrow V$ is étale and surjective, and (2) the two closed subschemes $R \times_V W$, $W \times_V W$ of $U \times U \times V$ are equal. Let $S \rightrightarrows V \rightarrow Y$ be an algebraic space. Then $\text{Hom}(Y, X)$ is defined as the kernel of $\text{Hom}(V, X) \rightrightarrows \text{Hom}(S, X)$. If Y is an affine scheme, this definition of $\text{Hom}(Y, X)$ is equivalent to the previous definition by virtue of the étale descent. Thus algebraic spaces form a category which contains the category of schemes. We can define the structure sheaf of an algebraic space and construct a cohomology theory. Many important notions and theorems for schemes can be generalized to those for algebraic spaces. Every algebraic space has a dense open subset that is an affine scheme. A group algebraic space is a group scheme (J. P. Murre). Suppose $k = \mathbb{C}$. If an algebraic group G operates on an algebraic k -scheme properly with a finite stabilizer group, the quotient space exists as an algebraic space. In this way H. Popp has shown the existence of moduli spaces of algebraic surfaces of general type as algebraic spaces [70, 71] (- 15 Algebraic Surfaces; also [54]). Moreover, every separated algebraic space X of finite type over \mathbb{C} carries a natural structure X^{an} of an analytic space. If there exists a proper modification morphism $f: X^{\text{an}} \rightarrow Y$ of a separated algebraic space X onto an analytic space Y , then Y carries a structure of an algebraic space and f becomes a morphism of algebraic spaces (Artin [2]). For any algebraically closed field k , Artin has introduced the notion of formal algebraic space and formal contraction and has obtained results similar to those for algebraic spaces [49]. An irreducible compact complex space X whose algebraic dimension (\rightarrow 72 Complex Manifolds F) is equal to $\dim X$ is called a

Moishezon space. As a corollary of the above theorem, any Moishezon space X carries a structure M of a compact algebraic space such that $X \cong M^{\text{an}}$.

X. Formal Schemes

Let A be a ring which we assume to be Noetherian, for simplicity, and I an ideal of A . Taking $\{I^n\}_{n>0}$ as a fundamental system of neighborhoods of 0, we can introduce a structure of a topological ring into A called **I -adic topology**. The completion of A with I -adic topology is isomorphic to the projective limit $\hat{A} = \varprojlim_{n>0} A/I^n$ (here A/I^n are regarded as discrete topological rings) and called the **completion of A along I** . If A is Noetherian, then \hat{A} is again Noetherian. There is a canonical continuous homomorphism $i: A \rightarrow \hat{A}$ whose kernel comprises the zero divisors a with $a \cdot 1 \in I$ (intersection theorem of Krull; → 284 Noetherian Rings B). If i is an isomorphism, we say A is **complete** with respect to I . The topology of \hat{A} is the \hat{I} -adic topology where $\hat{I} = i(I)\hat{A}$ and \hat{A} is complete with respect to \hat{I} . Take a Noetherian ring A complete with respect to I which we consider as an I -adic topological ring by identifying its completion along I with A . On $\mathfrak{X} = V(I) \subset \text{Spec}(A)$ we can define a sheaf of topological rings $\mathcal{O}_{\mathfrak{X}}$ by $\Gamma(\mathcal{D}(f), 0) = \varprojlim_{n>0} A_f/I^n A_f$ for $\mathcal{D}(f) = D(f) \cap \mathfrak{X}$ with $f \in A$. We call $(\mathfrak{X}, \mathcal{O}_{\mathfrak{X}})$ the **formal spectrum** of A and write $\text{Spf}(A)$. I is called a **defining ideal** of $\text{Spf}(A)$. A (locally Noetherian) **formal scheme** is by definition a topological local ringed space which is locally isomorphic to a formal spectrum (of a Noetherian ring). If we define morphisms between two formal schemes by those in the category of topological local ringed spaces, the formal schemes form a category.

For two formal spectra $\text{Spf}(A)$ and $\text{Spf}(B)$ with defining ideals I and J , respectively, the direct product $\text{Spf}(A) \times \text{Spf}(B)$ in the category of formal schemes is the formal spectrum of the completion of $A \otimes B$ along the image of $I \otimes B + A \otimes J$. Similarly, we can construct a fiber product of formal schemes. A formal scheme \mathfrak{X} is called **separated** if the image of the diagonal morphism $\Delta_{\mathfrak{X}}: \mathfrak{X} \rightarrow \mathfrak{X} \times \mathfrak{X}$ is closed (→ Section D).

For a Noetherian ring A with an ideal I , the formal spectrum $\text{Spf}(\hat{A})$ (with a defining ideal \hat{I}) is called the **completion of $\text{Spec}(A)$ along $P(I)$** . Similarly for a Noetherian scheme X and a closed subscheme X' we can define the **completion $X_{|X'}$ of X along X'** . Every completion of a separated scheme is separated. For a coherent sheaf F on X one can define its **completion $F_{|X'}$ along X'** , which is again coherent (under the assumption that X is locally Noetherian).

Thus we can develop a theory of formal schemes in a way similar to that of schemes, which we call **“formal geometry”** (for the more general definitions and further discussions see [2; 29, I, III; 303]). Roughly speaking, a function on $X_{|X'}$ is a formal Taylor series with respect to the direction normal to X' whose coefficients are regular functions on X' . The method of formal completion enables us to introduce “analytic” or “infinitesimal” methods in algebraic geometry. Among many important theorems, we state here the following two theorems. (1) **The fundamental theorem of proper mapping:** Let $f: X \rightarrow Y$ be a proper morphism of locally Noetherian schemes, Y' a closed subscheme of Y , and $X' = X \times_Y Y'$ the inverse image of Y' . Denote the respective completions of X and Y along X' and Y' by \hat{X} and \hat{Y} , respectively. We have the induced proper morphism of formal schemes $\hat{f}: \hat{X} \rightarrow \hat{Y}$. Then we have canonical isomorphisms, $(R^n f_*(F))_{|Y'} \cong R^n \hat{f}_*(F_{|X'})$, $n \geq 0$, for every coherent, \mathcal{O}_X -Module F on X . This theorem can be applied to prove **Zariski’s connectedness theorem:** for a proper morphism $f: X \rightarrow Y$ of locally Noetherian schemes with $f_*(\mathcal{O}_X) = \mathcal{O}_Y$, every fiber $f^{-1}(y)$ off is connected and non-empty for $y \in Y$. (2) We use the same notation as in (1) and assume, moreover, that $Y = \text{Spec}(A)$ for a Noetherian ring A , complete with respect to an ideal I , and $Y' = V(I)$. Then the correspondence $F \rightarrow F_{|X'}$ gives an equivalence between the category of coherent \mathcal{O}_X -Modules with proper support over Y and the category of coherent $\mathcal{O}_{\hat{X}}$ -Modules with proper support over \hat{Y} . This theorem plays an important role in the theory of \dagger deformations of algebraic varieties.

Y. Algebraic Vector Bundles

In this section a **vector bundle** is a locally free sheaf of finite constant rank (→ Section E). A quotient sheaf of a vector bundle is called a **quotient bundle** if it is a vector bundle. A subsheaf F of a vector bundle E is a **subbundle** when both F and E/F are vector bundles. A vector bundle is said to be **indecomposable** unless it is a direct sum of proper subbundles. Every vector bundle E on \mathbf{P}^1 is a direct sum of line bundles, that is, $E \cong \bigoplus \mathcal{O}(a_i)$ (Grothendieck). This property characterizes \mathbf{P}^1 in the category of nonsingular projective varieties. In fact, if X is a nonsingular projective variety with $\dim X > 0$ and $X \not\cong \mathbf{P}^1$ and if r is an integer with $r \geq \dim X$, there are stable (see below), a fortiori, indecomposable vector bundles on X of rank r (J. Simonis and M. Maruyama). Vector bundles are closely related to subschemes of the base variety. Let E be a

vector bundle of rank 2 on a nonsingular quasiprojective variety X and $\mathcal{O}(1)$ an ample invertible sheaf on X . For $n \gg 0$ and general $s \in H^0(X, E(n))$, $E(n)/s\mathcal{O}_X \cong L \otimes I$ with L a line bundle and I an ideal in \mathcal{O}_X which defines a smooth subscheme $Y = (s)_0$ of codimension 2. When X is, for example, \mathbf{P}^n ($n \geq 3$) or affine, the converse holds. Let Y be a subscheme of X purely of codimension 2 and locally of complete intersection. If $\omega_Y \cong A^4 \otimes \mathcal{O}_Y$ for an invertible sheaf M on X , there is a vector bundle E of rank 2 and $s \in H^0(X, E)$ with $(s)_0 = Y$. Moreover, E is decomposable if and only if Y is globally a complete intersection. An Abelian surface can be embedded in $\mathbf{P}^4(\mathbf{C})$, and hence we have an indecomposable vector bundle of rank 2 on $\mathbf{P}^4(\mathbf{C})$ [40]. For $n \geq 5$, we have no examples of indecomposable vector bundles of rank 2 on $\mathbf{P}^n(\mathbf{C})$ [36]. Every vector bundle on $A^n = k^n$ is trivial (D. Quillen [72], A. Suslin). From this and the fact stated above several results can be deduced; every nonsingular curve in A^3 is set-theoretically a complete intersection (L. Szpiro).

For a vector bundle E on a complete scheme, the following are equivalent: (i) for every coherent \mathcal{O}_X -module F , $F \otimes S^n(E)$ is generated by its global sections for $n \gg 0$, (ii) for every coherent sheaf F on X , $H^i(X, F \otimes S^n(E)) = 0$ for all $i > 0$ and $n \gg 0$, (iii) the tautological line bundle $\mathcal{O}(1)$ on $\mathbf{P}(E)$ is ample. A vector bundle having these properties is said to be **ample**. This is a generalization of the notion of ampleness of invertible sheaves (- Section E). The set of ample vector bundles is closed under several operations [34]. No good criterion such as Nakai-Moishezon's for the ampleness of invertible sheaves (- Section E) is yet known. The tangent bundle of a nonsingular complete variety X is ample if and only if $X \cong \mathbf{P}^n$ (- Section R; S. Mori [59]).

Let Y be a nonsingular projective variety over an algebraically closed field and $\mathcal{O}_Y(1)$ an ample invertible sheaf on Y . A coherent sheaf E on Y is said to be **stable** (or, **semistable**) with respect to $\mathcal{O}_Y(1)$ if E is torsion-free and if for every coherent subsheaf F of E with $0 \neq F \neq E$, $\chi(F(m))/r(F) < (\leq) \chi(E(m))/r(E)$ for all $m \gg 0$, where $r(*)$ denotes the rank. If E is stable (or semistable) and locally free, it is called a **stable** (resp. **semistable**) **vector bundle**. Let $f: X \rightarrow S$ be a smooth, projective, geometrically integral morphism and $\mathcal{O}_X(1)$ an f -ample invertible sheaf on X . For a numerical polynomial H and an S -scheme T , set $\Sigma_{X/S}^H(T) = \{E \mid E \text{ is a coherent sheaf on } X_T \text{ with the properties (a) and (b)}\} / \sim$; (a) E is T -flat, (b) for every geometric fiber X_t , $E_t = E|_{X_t}$ is stable with respect to $\mathcal{O}_{X_t}(1)$ and $\chi(E_t(m)) = H(m)$, and if

$E_1 \cong E_2 \otimes f_T^*(L)$ with L an invertible sheaf on T , $E_1 \sim E_2$. $\Sigma_{X/S}^H$ is a contravariant functor of the category of locally Noetherian S -schemes to the category of sets. If S is of finite type over a universally Japanese ring A , $\Sigma_{X/S}^H$ has a coarse moduli scheme $M_{X/S}(H)$ (- Section W) and it is locally of finite type over S (D. Mumford, Seshadri, D. Gieseker, Maruyama [23, 51]). The set of the classes of semistable sheaves under a suitable equivalence relation (S -equivalence) on geometric fibers of X over S also has a coarse moduli scheme $M_{X/S}(H)$ and $M_{\text{ss}}(H)$ is its open subscheme. Moreover, it is known that $M_{\text{ss}}(H)$ is projective over S in some cases, for example, when A is a field of characteristic zero. When $\dim X/S = 1$, $X \cong \mathbf{P}^2$ or \mathbf{P}^3 , the structures of $M_{\text{ss}}(H)$ have been extensively studied [11, 69]. Theories of vector bundles are used in theoretical physics [9].

Z. Torus Embeddings

Over an algebraically closed field k , a **torus embedding**, or a **toric variety**, is a normal scheme X locally of finite type over k on which an algebraic torus T acts with a dense open orbit isomorphic to T . Such X 's, as well as many of their algebrogeometric properties, can be described very simply in terms of cones in real affine spaces, as Demazure [21] first saw in the nonsingular case in connection with algebraic subgroups of the Cremona transformation group, and then as D. Mumford et al. [47] as well as K. Miyake and T. Oda [67] saw in the general case immediately after H. Sumihiro [SS] proved a basic theorem on linear algebraic group actions.

The group N , written additively, of one-parameter subgroups of an r -dimensional algebraic torus T is a free Abelian group of rank r . A **convex rational polyhedral cone** σ in $N_{\mathbf{R}} = \mathbf{R} \otimes_{\mathbf{Z}} N$ with 0 as the vertex is the set of nonnegative linear combinations of a finite number of elements of N such that $\sigma \cap (-\sigma) = \{0\}$. A subset τ of σ is called a **face**, and is denoted by $\tau < \sigma$, if there exists a linear functional m on $N_{\mathbf{R}}$ having nonnegative values on σ and $\tau = \{y \in \sigma \mid m(y) = 0\}$. A **fan** Δ in $N_{\mathbf{R}}$ is a collection of such σ 's satisfying the conditions (i) $\Delta \ni \sigma, \sigma > \tau \Rightarrow \Delta \ni \tau$ and (ii) $\Delta \ni \sigma, \sigma' \Rightarrow \sigma > \sigma \cap \sigma' < \sigma'$.

Each $\sigma \in \Delta$ gives rise to an affine torus embedding U_{σ} as follows: Let M be the \mathbf{Z} -module dual to N ; hence A^4 is the group of characters of T . For $\sigma \in \Delta$, the set $M \cap \sigma = \{m \in M \mid m(y) \geq 0 \text{ for all } y \in \sigma\}$ is seen to be a finitely generated additive subsemigroup of M containing 0 and generating M as a group. U_{σ} is then the spectrum of the semigroup algebra $k[M \cap \sigma]$.

Thanks to the condition (ii) above, the U_σ 's can be naturally pasted together to produce a torus embedding $X = \bigcup_{\sigma \in \Delta} U_\sigma$.

Every torus embedding is obtained in this way. Equivariant dominant morphisms between torus embeddings can be described in terms of \mathbb{Z} -linear maps between N 's which send a fan to another.

Many algebrogeometric properties of X can be described in terms of A , e.g., X is nonsingular if and only if every $\sigma \in A$ can be spanned by a part of a \mathbb{Z} -basis of N . X is complete (proper over k) if and only if the union of σ 's in A coincides with N . X has at worst cyclic quotient singularities if and only if every $\sigma \in A$ is simplicial, i.e., if it is spanned by R -linearly independent elements of N . An equivariant resolution of singularities of X exists and is obtained by a suitable subdivision of A (Mumford et al. [47]). X has only rational singularities, hence is Cohen-Macaulay (M. Hochster [38] and Mumford et al. [47]). The set of T -orbits in X is in one-to-one correspondence with $\sigma \in \Delta$. A reduced T -invariant subscheme Y of X is the union of T -orbits, and hence corresponds to a subset Σ of A . M.-N. Ishida [44] determined when Y is Cohen-Macaulay or Gorenstein in terms of the combinatorics of Σ .

A T -invariant Cartier divisor D on X corresponds to a support function h , which is a continuous R -valued function on the union $\bigcup_{\sigma \in \Delta} \sigma$ which is (i) positively homogeneous, i.e., $h(\lambda y) = \lambda h(y)$ for $\lambda > 0$ and $y \in \bigcup_{\sigma \in \Delta} \sigma$, (ii) \mathbb{Z} -valued on $N \cap (\bigcup_{\sigma \in \Delta} \sigma)$ and (iii) linear when restricted to each $\sigma \in \Delta$. Various properties of the invertible sheaf $\mathcal{O}_X(D)$ (\rightarrow Section E) can be described in terms of h . For instance, when X is complete, $\mathcal{O}_X(D)$ is t ample if and only if h is upper convex, i.e., $h(y) + h(y') \leq h(y + y')$, and moreover, A is the coarsest fan with the property (iii) above. The cohomology of $\mathcal{O}_X(D)$ can be calculated by means of h (Demazure [21] and V. Danilov [17]).

A support function h , on the other hand, gives rise to a convex polyhedron in M_R with vertices in M . In this way, certain aspects of the geometry of convex sets can be thought of as a part of the theory of projective varieties (R. Stanley [87], B. Teissier [89]).

Mumford et al. [47] introduced a more general concept of **toroidal embedding**: A normal algebraic variety Y and a nonsingular Zariski open subset U such that $Y \supset U$ is formally isomorphic at each point to a torus embedding $X \supset T$. This concept has been used very effectively to prove important theorems: (1) Systematic nice compactifications of arithmetic quotients of bounded symmetric domains (D. Mumford et al. [7], I. Satake [77]). Y. Nami-kawa [65] worked out the details in the case

of the Siegel upper planes, improving their earlier Satake compactifications studied by I. Satake, W. Baily, J.-I. Igusa, and A. Borel [12].

(2) **Semistable reduction theorem** (Mumford et al. [47]). Let $f: V \rightarrow C$ be a flat morphism from a nonsingular variety V to a nonsingular curve C in characteristic 0. After a finite base extension $C' \rightarrow C$ and a modification $V' \rightarrow V \times_C C'$, we can get $f': V' \rightarrow C'$ whose singular fibers are reduced with only nonsingular components crossing normally with each other (\rightarrow Section L). Without the reducedness requirement, the existence is derived from a result of Hironaka [37].

AA. Etale Topology

Let \mathcal{S} be a category. We say that a **Grothendieck topology** on Y is given if, for each $S \in \text{Ob}(\mathcal{S})$, families of morphisms (**covering families** of S) are given and satisfy the following conditions: (1) If $\varphi: T \rightarrow S$ is an isomorphism, $\{\varphi: T \rightarrow S\}$ is a covering family, (2) if $\{\varphi_i: R_i \rightarrow S\}_{i \in I}$ is a covering family, for any morphism $\varphi': S' \rightarrow S$, the fiber product $R'_i = R_i \times_S S'$ exists and the induced family $\{\varphi'_i: R'_i \rightarrow S'\}$ is a covering family of S' , (3) if $\{\varphi_i: R_i \rightarrow S\}_{i \in I}$ is a covering family and if for each $i \in I$, $\{s_{i,a}: S_{i,a} \rightarrow R_i\}_{a \in A_i}$ is a covering family, then $\{\varphi_i^0 s_{i,a}: S_{i,a} \rightarrow S\}_{i \in I, a \in A_i}$ is a covering family. In general, it is more convenient to use a notion of sieves to define a Grothendieck topology [30 (SGA 4)]. A category with Grothendieck topology is called a **site**. Let X be a scheme and Et/X the category whose objects consists of schemes étale over X . If we choose a family of morphisms $\{\phi_i: Y_i \rightarrow Y\}_{i \in I}$, with $\bigcup_{i \in I} \phi_i(Y_i) = Y$, as a covering family of $Y \in \text{Ob}(\text{Et}/X)$, this defines the **étale topology** on X and the **étale site** $X_{\text{ét}}$. Similarly, one can define the **Zariski site** X_{Zar} (resp. the **flat site** X_{fl}), using open immersions (resp. flat morphisms locally of finite type). A **presheaf** on a **site** Y is a contravariant functor from Y to (Set) . A presheaf F is a **sheaf** if, for any covering family $\{\phi_i: R_i \rightarrow R\}_{i \in I}$, $F(R) \rightarrow \prod_{i \in I} F(R_i) \rightrightarrows \prod_{i,j \in I} F(R_i \times_R R_j)$ is exact. If Y has a final object X , the functor $F \mapsto F(X)$ is left exact on the category of Abelian sheaves on \mathcal{S} and the cohomology groups $H^*(X, F)$ are defined as the right derived functors. Using covering families, one can define the Čech cohomology $\check{H}^*(X, F)$ as usual. In the following, a sheaf means an Abelian sheaf. For a t geometric point $i: x \rightarrow X$ of a scheme, an étale neighborhood U of x consists of an étale morphism $f: U \rightarrow X$ and a morphism $j: x \rightarrow U$ such that $f \circ j = i$. For a sheaf F on $X_{\text{ét}}$, the stalk F_x at a geometric point x is defined by $F_x =$

$\lim \Gamma(U, F)$, U being the étale neighborhoods of x . On $X_{\text{ét}}$ the sheaf \mathbf{G}_m is defined by $G_i(S) = \Gamma(S, \mathcal{O}_S)^*$ for each $S \in \text{Ob}(\text{Et}/X)$. The kernel of the n th power homomorphism $\mathbf{G}_m \xrightarrow{f_n} \mathbf{G}_m$ is denoted by μ_n . The sequence $1 \rightarrow \mu_n \rightarrow \mathbf{G}_m \xrightarrow{f_n} \mathbf{G}_m \rightarrow 1$ is exact in $X_{\text{ét}}$ if p is prime to residual characteristics of X , but not necessarily exact in X_{zar} . There are canonical isomorphisms $\text{Pic}(X) = H^1(X_{\text{zar}}, \mathcal{O}_X^*) \simeq H^1(X_{\text{ét}}, \mathbf{G}_m) \simeq H^1(X_{\text{fl}}, G_i)$. For $X = \text{Spec}(k)$, k a field, the étale cohomology theory of sheaves on $X_{\text{ét}}$ is equivalent to the \dagger Galois cohomology theory over k ; hence $H^i(X_{\text{ét}}, F) \simeq H^i(G, F_k)$, where $G = \text{Gal}(k/\bar{k})$, \bar{k} the \dagger separable closure of k and F_k is the stalk of F at the geometric point \otimes_k of X . For a morphism $f: X \rightarrow Y$ of schemes and a sheaf F on $X_{\text{ét}}$, the direct image sheaf $f_* F$ of F is defined by $f_* F(S) = F(X \times_Y S)$, $S \in \text{Ob}(\text{Et}/Y)$, and higher direct-image sheaves $R^i f_* F$ are defined by the right-derived functor. Let X be a separated scheme of finite type over a field k . By a theorem of Nagata [63], there exists a scheme X proper over k and an open immersion $j: X \rightarrow \bar{X}$. For a torsion sheaf F on $X_{\text{ét}}$, let $j_! F$ be a sheaf on $\bar{X}_{\text{ét}}$ extended by 0 outside X . The cohomology with compact support $H_c^q(X_{\text{ét}}, F)$ is defined by $H_c^q(X_{\text{ét}}, F) = H^q(\bar{X}_{\text{ét}}, j_! F)$. This is independent of compactifications. Similarly, for a separated morphism $f: X \rightarrow S$ of finite type of schemes and a torsion sheaf F on $X_{\text{ét}}$, one can define higher direct image sheaf with compact support $R^q f_! F$. For the étale topology, torsion sheaves are important. All torsion sheaves on $X_{\text{ét}}$ on a Noetherian scheme X are inductive limits of constructible sheaves. A sheaf F on $X_{\text{ét}}$ is called **locally constructible (constant)** if F is represented by an étale covering of X . A sheaf F on $X_{\text{ét}}$ is called **constructible** if there exists a finite surjective family of subscheme X_i of X such that the restriction of F to X_i is locally constructible (constant). This is equivalent to saying that every irreducible closed subscheme Z of X contains a nonempty open subscheme U such that the restriction F to U is locally constant and has finite stalks, i.e., there is a covering family $\{\phi_i: U_i \rightarrow U\}$ such that $F|_{U_i}$ is constant and the stalk F_x is a finite Abelian group for all geometric points x of U . The cohomology of a torsion sheaf or a constructible sheaf has properties similar to those of the classical cohomology. Let $f: X \rightarrow S$ be a proper morphism and F a torsion sheaf on $X_{\text{ét}}$. Then the stalk $R^q f_* F$ at a geometric point s of S is isomorphic to $H^q(X_{s, \text{ét}}, F)$, where $X_s = X \times_S \text{Spec } k(s)$. If f is a separated morphism of finite type of Noetherian schemes, a similar fact holds for the cohomology with compact support. Moreover, if F is a constructible sheaf, $R^q f_! F$ is also constructible (**finiteness theorem**). For an affine scheme X of

finite type over a separably closed field and F a torsion sheaf on $X_{\text{ét}}$, one has $H^q(X_{\text{ét}}, F) = 0$ for $q > \dim X$. Let $f: X \rightarrow S$ be a separated morphism of schemes of finite type on C , F a torsion sheaf on $X_{\text{ét}}$. One has a canonical isomorphism $(R^q f_! F)^{an} \simeq R^q f_!^{an} F^{an}$. In particular, if X is proper over C , one has $H^q(X_{\text{ét}}, \mathbf{Z}/(n)) \simeq H^q(X^{an}, \mathbf{Z}/(n))$. On the other hand, for a nonsingular complete curve C over an algebraically closed field, one always has $H^1(C_{\text{ét}}, \mathbf{Z}) = 0$. For a geometric point s of S the **strict localization** of S at s is the ring $\mathcal{O}_{S, s} = \varinjlim \Gamma(U, \mathcal{O}_U)$, U being étale neighborhoods of s . A geometric point t of S is called a **generalization** of s if t is defined by an algebraic closure of the residue field of a point of $\text{Spec } \mathcal{O}_{S, s}$. In this case, s is called a **specialization** of t . If $f: X \rightarrow S$ is a proper smooth morphism the sheaf $R^q f_* \mathbf{Z}/(n)$ is locally constructible (constant), and if s is a specialization of t one can define a **cospecialization** of $H^q(X_{t, \text{ét}}, \mathbf{Z}/(n))$ to $H^q(X_{s, \text{ét}}, \mathbf{Z}/(n))$ which is bijective. Let X be a scheme of finite type over a field k . If $l \nmid \# \text{ch}(k)$ is prime, one can define the l -adic cohomology $H^q(X, \mathbf{Q}_l) = \varinjlim_n H^q(X, \mathbf{Z}/(l^n)) \otimes_{\mathbf{Z}_l} \mathbf{Q}_l$. The l -adic cohomology has as many good properties as the classical cohomology (\rightarrow Section U).

References

- [1] A. Altman and S. L. Kleiman, Introduction to Grothendieck duality theory, Lecture notes in math. 146, Springer, 1970.
- [2] M. Artin, Algebraization of formal moduli; I, Global analysis, Univ. Tokyo Press and Princeton Univ. Press, 1969, 21-71; II, Existence of Modifications, Ann. Math., (2) 91 (1970), 88-135.
- [3] M. Artin, Algebraic approximation of structures over complete local rings, Publ. Math. Inst. HES, 36 (1969), 233-58.
- [4] M. Artin, Théorèmes de représentabilité pour les espaces algébriques, Les Presses de l'Université de Montréal, 44, 1973.
- [5] M. Artin and B. Mazur, Etale homotopy, Lecture notes in math. 100, Springer, 1969.
- [6] M. Artin and D. Mumford, Some elementary examples of unirational varieties which are not rational, Proc. London Math. Soc., (3) 25 (1972), 755-95.
- [7] A. Ash, D. Mumford, M. Rapoport, and Y. Tai, Smooth compactification of locally symmetric varieties, Math.-Sci. Press, 1975.
- [8] M. F. Atiyah, Vector bundles over an elliptic curve, Proc. London Math. Soc., (3) 7 (1957), 414-452.
- [9] M. F. Atiyah and R. S. Ward, Instantons and algebraic geometry, Comm. Math. Phys., 55 (1977), 117-124.

- [10] P. Berthelot, Cohomologie cristalline des schémas de caractéristique $p > 0$, Lecture notes in math. 407, Springer, 1974.
- [11] W. Barth, Some properties of stable rank-2 vector bundles on P_n , Math. Ann., 226 (1977), 1255-150.
- [12] W. L. Baily and A. Borel, Compactification of arithmetic quotients of bounded symmetric domains, Ann. Math., (2) 84 (1966), 442-528.
- [13] S. Bloch, D. Lieberman, and A. Kas, Zero cycles on surfaces with $p_g = 0$, Compositio Math., 33 (1976), 135-145.
- [14] E. Brieskorn, Rationale Singularitäten Komplexer Flächen, Inventiones Math., 4 (1968), 336-358.
- [15] Séminaire C. Chevalley, Anneaux de Chow et applications, Ecole Norm. Sup., 1958.
- [16] C. H. Clemens and P. A. Griffiths, The intermediate Jacobian of the cubic threefold, Ann. Math., (2) 95 (1972), 281-356.
- [17] V. I. Danilov, The geometry of toric varieties, Russian Math. Surveys, 33 : 2, (1978), 977-154.
- [18] P. Deligne, Théorie de Hodge: I, Actes Congr. Intern. Math., 1970, Nice, Gauthier-Villars, 1971, vol. 1, 425-430; II, Publ. Math. Inst. HES, 40 (1971), 5-57; III, 44 (1974), 5577.
- [19] P. Deligne, La conjecture de Weil I, Publ. Math. Inst. HES, 43 (1974), 273-307.
- [20] P. Deligne, Le groupe fondamental du complément d'une courbe plane n'ayant que des points doubles ordinaires est abélien (d'après W. Fulton), Séminaire Bourbaki, exp. 543, Nov. 1979, Lecture notes in math. 842, Springer, 1980.
- [21] M. Demazure, Sous-groupes algébriques de rang maximum du groupe de Cremona, Ann. Sci. Ecole Norm. Sup., 3 (1970), 507-588.
- [22] W. Fulton and J. Hansen, A connectedness theorem for projective varieties, with applications to intersections and singularities of mappings, Ann. Math., (2) 110 (1979), 159-166.
- [23] D. Gieseker, On the moduli of vector bundles on an algebraic surface, Ann. Math., (2) 106 (1977), 45-60.
- [24] P. A. Griffiths, Periods of integrals on algebraic manifolds; summary of main results and discussion of open problems, Bull. Amer. Math. Soc., 76 (1970), 228-296.
- [25] P. A. Griffiths, On the periods of certain rational integrals I, II, Ann. Math., (2) 90 (1969), 460-495, 496-541.
- [26] P. A. Griffiths, Periods of integrals on algebraic manifolds III, Publ. Math. Inst. HES, 38 (1970), 1255-180.
- [27] P. A. Griffiths and W. Schmid, Recent developments in Hodge theory, Discrete Subgroups of Lie Groups and Application to Moduli, Oxford Univ. Press, 1975, 3 1-127.
- [28] P. A. Griffiths and J. Harris, Principles of algebraic geometry, Wiley, 1978.
- [29] A. Grothendieck, Eléments de géométrie algébrique, Publ. Math. Inst. HES: 1, 4 (1960) (Springer, 1971); II, 8 (1961); III, 11 (1961), 17 (1963); IV, 20 (1964), 24 (1965), 28 (1966), 32 (1967).
- [30] Cited as SGA. A. Grothendieck and others, Séminaire de géométrie algébrique: SGA 1, Revêtements étales et groupe fondamental, Lecture notes in math. 224, Springer, 1971; SGA 2, Cohomologie locale des faisceaux et théorèmes de Lefschetz locaux et globaux, North-Holland, 1962; SGA 3 (with M. Demazure), Schémas en groupes I, II, III, Lecture notes in math. 15 1, 152, 153, Springer, 1970; SGA 4 (with M. Artin and J. L. Verdier), Théorie des topos et cohomologie étale des schémas, Lecture notes in math. 269, 270, 305, Springer, 1972-1973; SGA 5, Cohomologie l -adique et fonctions L ; SGA 6, Théorie des intersections et théorème de Riemann-Roch, Lecture notes in math. 225, Springer, 1971; SGA 7, Groupes de monodromie en géométrie algébrique, pt. I, Lecture notes in math. 288; pt. II (by P. Deligne and N. Katz), 340, Springer, 1972-1973.
- [31] B. Grünbaum, Polytopes, graphs and complexes, Bull. Amer. Math. Soc., 76 (1970), 1131-1201.
- [32] W. J. Haboush, Reductive groups are geometrically reductive, Ann. Math., (2) 102 (1975), 677-83.
- [33] R. Hartshorne, Residues and duality, Lecture notes in math. 20, Springer, 1966.
- [34] R. Hartshorne, Ample subvarieties of algebraic varieties, Lecture notes in math. 156, Springer, 1970.
- [35] R. Hartshorne, Algebraic geometry, Graduate texts in math. 52, Springer, 1977.
- [36] R. Hartshorne, Algebraic vector bundles on projective spaces: A problem list, Topology, 18 (1979), 117-128.
- [37] H. Hironaka, Resolution of singularities of an algebraic variety over a field of characteristic zero, Ann. Math., (2) 79 (1964), 109-326.
- [38] M. Hochster, Rings of invariants of tori, Cohen-Macaulay rings generated by monomials and polytopes, Ann. Math., (2) 96, (1972), 318-337.
- [39] W. V. D. Hodge and M. F. Atiyah, Integrals of the second kind on an algebraic variety, Ann. Math., (2) 62 (1955), 56-91.
- [40] G. Horrocks and D. Mumford, A rank-2 vector bundle on P^4 with 15,000 symmetries, Topology, 12 (1973), 63-81.
- [41] J. Igusa, Arithmetic variety of moduli for genus two, Ann. Math., (2) 72 (1960), 612-649.
- [42] J. Igusa, Betti and Picard numbers of

abstract algebraic surfaces, *Proc. Nat. Acad. Sci. US*, 46 (1960), 7244726.

[43] H. Inose and M. Mizukami, Rational equivalence of O -cycles on some surfaces of general type with $p_g = 0$, *Math. Ann.*, 244 (1979), 205-217.

[44] M.-N. Ishida, Torus embeddings and dualizing complexes, *Tôhoku Math. J.*, 32, (1980), 111-146.

[45] V. A. Iskovskii and Yu. I. Manin, Three-dimensional quartics and counterexamples to the Lüroth problem, *Math. USSR-Sb.*, 15 (1971), 141-166. (Original in Russian, 1971.)

[46] S. Lojasiewicz, Triangulation of semianalytic sets, *Ann. Scuola Norm. Sup. Pisa*, 18 (1964), 449-474.

[47] G. Kempf, F. Knudsen, D. Mumford, and B. Saint-Donat, *Toroidal embeddings I*, Lecture notes in math. 339, Springer, 1973.

[48] S. L. Kleiman, Toward a numerical theory of ampleness, *Ann. Math.*, 84 (1966), 293-344.

[49] D. Knutson, *Algebraic spaces*, Lecture notes in math. 203, Springer, 1971.

[SO] K. Kodaira, *Collected works I-III*, Iwanami, 1975.

[51] M. Maruyama, Moduli of stable sheaves, I, II, *J. Math. Kyoto Univ.*, 17 (1977), 91-126 and 18 (1978), 557-614.

[52] T. Matsusaka, *Theory of Q -varieties*, Publ. Math. Soc. Japan, no. 8, Tokyo, 1965.

[53] T. Matsusaka, Polarized varieties with a given Hilbert polynomial, *Amer. J. Math.*, 94 (1972), 1027-1077.

[54] T. Matsusaka and D. Mumford, Two fundamental theorems on deformations of polarized varieties, *Amer. J. Math.*, 86 (1964), 668-684.

[55] M. Miyanishi, On curves on rational and unirational surfaces, *Tata lecture notes* 60, Springer, 1978.

[56] B. G. Moishezon, A criterion for projectivity of complete algebraic abstract varieties, *Amer. Math. Soc. Transl.*, 63 (1967), 1-50.

[57] B. G. Moishezon, Complex surfaces and connected sums of complex projective planes, *Lecture notes in math.* 603, Springer, 1977.

[SS] S. Mori, On a generalization of complete intersections, *J. Math. Kyoto Univ.*, 15 (1975), 619-646.

[59] S. Mori, Projective manifolds with ample tangent bundles, *Ann. Math.*, (2) 110 (1979), 593-606.

[60] D. Mumford, *Geometric invariant theory*, Springer, 1965; second edition, 1982.

[61] D. Mumford, Pathologies of modular algebraic surfaces, *Amer. J. Math.*, 83 (1961), 339-342; Further pathologies in algebraic geometry, *ibid.*, 84 (1962), 642-648; Pathologies III, *ibid.*, 89 (1967), 944-104.

[62] D. Mumford, Stability of projective

varieties, *Enseignement Math.*, 23 (1977), 39-110.

[63] M. Nagata, Imbedding of an abstract variety in a complete variety, *J. Math. Kyoto Univ.*, 2 (1962), 1-10.

[64] Y. Nakai, A criterion of an ample sheaf on a projective scheme, *Amer. J. Math.*, 85 (1963), 14-26.

[65] Y. Namikawa, A new compactification of the Siegel space and the degeneration of Abelian varieties, I, II, *Math. Ann.*, 221 (1976), 97-141 and 201-241.

[66] A. Néron, Problèmes arithmétiques et géométriques attachés à la notion du rang d'une courbe algébrique, *Bull. Soc. Math. France*, 80 (1952), 101-166.

[67] T. Oda, On torus embeddings and applications (based on joint work with K. Miyake), *Tata lecture notes* 57, Springer, 1978.

[68] T. Oda and C. S. Seshadri, Compactifications of the generalized Jacobian variety, *Trans. Amer. Math. Soc.*, 253 (1979), 1-90.

[69] C. Okonek, M. Schneider, and H. Spindler, *Vector bundles on complex projective spaces*, Progress in math. 3, Birkhäuser, 1980.

[70] H. Popp, On moduli of algebraic varieties II, *Compositio Math.* 28 (1974), 51-81; III, Fine moduli spaces, *ibid.*, 31 (1975), 237-258.

[71] H. Popp, Moduli theory and classification theory of algebraic varieties, *Lecture notes in math.* 620, Springer, 1977.

[72] D. Quillen, Projective modules over polynomial rings, *Inventiones Math.*, 36 (1976), 167-171.

[73] M. A. Reid, Canonical 3-folds, *Journées de Géométrie Algébrique*, July 1979, 273-310, Sijthoff and Noordhoff, 1980.

[74] A. A. Rojzman, Rational equivalence of zero cycles, *Math. USSR-Sb.*, 18 (1972), 571-588. (Original in Russian, 1972.)

[75] A. A. Rojzman, The torsion of the group of O -cycles modulo rational equivalence, *Ann. Math.*, (2) 111 (1980), 553-569.

[76] L. Roth, *Algebraic threefolds*, Springer, 1955.

[77] I. Satake, On the arithmetic of tube domains, *Bull. Amer. Math. Soc.*, 79 (1973), 1076-1094.

[78] M. Schlessinger, Functors of Artin rings, *Trans. Amer. Math. Soc.*, 130 (1968), 208-222.

[79] C. S. Seshadri, Quotient spaces modulo reductive algebraic groups, *Ann. Math.*, (2) 95 (1972), 511-556.

[SO] J.-P. Serre, Un théorème de dualité, *Comm. Math. Helv.*, 29 (1955), 9-26.

[81] J.-P. Serre, Faisceaux algébriques cohérents, *Ann. Math.*, (2) 61 (1955), 197-278.

[82] J.-P. Serre, Géométrie algébrique et géométrie analytique, *Ann. Inst. Fourier*, 6 (1956), 1-42.

[83] J.-P. Serre, Exemple de variété projectives

- conjuguées non homéomorphes, C. R. Acad. Sci. Paris, 258 (1964), 419444196.
- [84] I. R. Shafarevich, Basic algebraic geometry, Springer, 1974.
- [SS] G. Shimura, Moduli and fibre systems of Abelian varieties, Ann. Math., (2) 83 (1966), 294-338.
- [86] G. Shimura, Algebraic varieties without deformation and the Chow variety, J. Math. Soc. Japan, 20 (1968), 336-341.
- [87] R. Stanley, The number of faces of a simplicial convex polytope, Advances in math., 35 (1980), 236-238.
- [88] H. Sumihiro, Equivariant completion I, J. Math. Kyoto Univ. 14 (1974), 1-28; II, 15 (1975), 573-605.
- [89] B. Teissier, Variétés toriques et polytopes, Sémin. Bourbaki, 1980/81, exp. 565, Lecture notes in math. 901, Springer, 1981.
- [90] K. Ueno, Classification theory of algebraic varieties and compact complex spaces, Lecture notes in math. 439, Springer, 1975.
- [91] H. Umemura, Sur les sous-groupes algébriques primitifs du groupe de Cremona a trois variables, Nagoya Math. J., 79 (1980), 477-67.
- [92] A. Weil, Foundations of algebraic geometry, Amer. Math. Soc. Colloq. Publ., second edition, 1962.
- [93] O. Zariski, Collected papers, I, II, III, IV, MIT Press, 1972, 1973, 1978, 1979.

17 (X1.11) Algebroidal Functions

A. General Remarks

If an analytic function f satisfies an \dagger irreducible algebraic equation

$$A_0(z)f^k + A_1(z)f^{k-1} + \dots + A_k(z) = 0 \quad (1)$$

with single-valued \dagger meromorphic functions $A_j(z)$ in a domain G in the complex z -plane, then f is called a **k -valued algebroidal function** in G . With no loss of generality, we can assume that there is no common zero among the $A_j(z)$ and that all the $A_j(z)$ are \dagger holomorphic in G . When $k = 1$, the solution of (1) is a single-valued meromorphic function in G . If all the $A_j(z)$ are polynomials, then f is an \dagger algebraic function. Thus algebroidal functions can be regarded as extensions of single-valued to multiple-valued functions and also as extensions of algebraic to \dagger transcendental functions. Since (1) is irreducible, its discriminant $D(z)$ does not vanish identically. For all the points a satisfying $D(a) \neq 0$, $A_0(a) \neq 0$, $a \in G$, equation (1) determines k holomorphic function ele-

ments $f_1(z), f_k(z)$ in a suitable neighborhood of a that determine the analytic function f . They can be \dagger prolonged analytically in G in the wider sense. At any point satisfying $A_0(z) = 0$, at least one element has a pole; and at any point satisfying $D(z) = 0$, there may appear ramified elements. Therefore an algebroidal function can be defined as a finitely multiple-valued analytic function in G with the exception of poles and \dagger algebraic branch points. Every algebroidal function $f(z)$ determines a \dagger Riemann surface, which may be considered a \dagger covering surface \mathfrak{Z} of G . This surface \mathfrak{Z} is a k -sheeted \dagger covering surface over G with no singular point except for algebraic branch points. Also, f reduces to a single-valued meromorphic function on \mathfrak{Z} , and all the function elements over a point z are different. A k -valued algebroidal function can also be characterized by this property.

These two (equivalent) definitions of algebroidal functions give rise to two distinct methods of studying these functions. In adopting the first definition, we can make use of results in the single-valued case, as did G. Rémoundos and G. Valiron. When the latter definition is adopted, we can use several methods that are also applicable in the single-valued case, as did H. Selberg and E. Ullrich.

Research on algebroidal functions has been carried out mainly for the case where G is the finite plane $|z| < \infty$ or the unit disk $|z| < 1$. Almost all the known results for algebroidal functions are extensions of those on single-valued functions, but several results particularly relevant to algebroidal functions have been discovered. The existence of branch points makes it difficult to investigate algebroidal functions in some cases.

B. Absolute Value

Among several results on the absolute values of algebroidal functions, the \dagger maximum principle, one of the basic principles, holds on the Riemann surface \mathfrak{Z} . The following relation holds among the $A_j(z)$ and $|f_v(z)|$. Assume that (1) has the form

$$f^k + A_1(z)f^{k-1} + \dots + A_k(z) = 0. \quad (1')$$

Then $\log(1 + A(z))/\log(1 + F(z))$ is bounded, where $A(z) = \max |A_j(z)|$ and $F(z) = \max |f_v(z)|$.

An algebroidal function that has no pole in $|z| < \infty$ is called an **entire algebroidal function**. The successive derivatives of an entire algebroidal function may have poles at every branch point, which is a departure from the case of single-valued integral functions. An algebroidal function defined by (1) or (1') with entire functions $A_j(z)$ and zero-free $A_0(z)$ is

entire. Dividing both sides of (1) by $A(z)$, we obtain equation (1'), and all its coefficients are entire. In equation (1'), the order, type, and class of an entire algebroidal function coincide with those of the largest $A_j(z)$ (429 Transcendental Entire Functions).

For an entire algebroidal function of order less than $1/2$, $F(z)$ tends to infinity uniformly along a sequence of concentric circles $|z|=r_n$ ($r_n \rightarrow \infty$). However, it can be shown that not all the branches of $f(z)$ necessarily tend to infinity on the Riemann surface \mathfrak{Z} ; in this sense, Wiman's theorem does not remain true. But, using $\min_{|z|=r} \max_v |f_v(z)|$, one can obtain some extensions of the generalized Wiman theorem.

C. Picard's Theorem and Its Extension

Rémoundos first extended Picard's theorem and Borel's theorem to an algebroidal function. Every k -valued transcendental algebroidal function in the finite plane takes on every value infinitely often with at most $2k$ exceptional values. There are examples where $2k$ values are actually omitted. Hence the theorem is the best one possible in this sense. T. Varopoulos (*Bull. Soc. Math. France*, 53 (1925), 23-34) introduced the **degeneracy index**

$$\lambda = \dim \{ (c_0, c_1, \dots, c_k) \in C^{k+1} \mid c_0 A_0 + c_1 A_1 + \dots + c_k A_k = 0 \}$$

($0 \leq \lambda \leq k-1$) into equation (1) and showed that the number of Picard's exceptional values off is at most $k + \lambda + 1$. There is no single-valued meromorphic function with $A > 0$, so this result is relevant only for algebroidal functions. Extending this idea further, J. Dufrénoy and others obtained more precise results. The situation is the same for the Borel exceptional values. That is, the convergence exponent of $f(z) = w = 0$ coincides with the order off except for at most $2k$ values, for which the convergence exponents off are less than the order off (E. Borel). There are at most $2k$ polynomials $P(z)$ for which $f(z) = P(z) = 0$ has at most a finite number of roots (Borel), and furthermore, using the degeneracy index, we can give more precise results.

Selberg was the first to extend the Nevanlinna theory of meromorphic functions to algebroidal functions (272 Meromorphic Functions). Almost simultaneously, Valiron obtained the same results starting from the coefficients of (1); then Ullrich improved the results by considering the effect of branch points of \mathfrak{Z} .

Let \mathfrak{Z}_r be the part of \mathfrak{Z} over $|z| < r$, let $n(r, w)$ be the number of roots of $f(z) = w = 0$ in \mathfrak{Z}_r , and let $n(r, \mathfrak{Z})$ be the number of branch points

of \mathfrak{Z}_r . With these notations we write

$$N(r, w) = \frac{1}{k} \int_0^r (n(t, w) - n(0, w)) \frac{dt}{t} + \frac{n(0, w)}{k} \log r,$$

$$m(r, w)$$

$$= \frac{1}{2k\pi} \int_{|z|=r} \log^+ \frac{1}{|f(re^{i\varphi}) - w|} d\varphi, \quad w \neq \infty,$$

$$T(r, w) = m(r, w) + N(r, w).$$

Let $T(r, f)$ be the logarithmic integral of the spherical area of the image of \mathfrak{Z}_r under $w = f(z)$,

$$T(r, f) = \frac{1}{k} \int_0^r \frac{dt}{t} \int_{\mathfrak{Z}_t} \frac{|f'(te^{i\varphi})|^2}{(1 + |f(te^{i\varphi})|^2)^2} t d\varphi,$$

and let $N(r, \mathfrak{Z})$ be the logarithmic integral of $n(r, \mathfrak{Z})$. Then we have $T(r, w) = T(r, f) + O(1)$ and the **ramification theorem**:

$$N(r, \mathfrak{Z}) < (2k-2)T(r, f) + O(1).$$

Also, $T(r, f) = O(\log r)$ holds if and only if f is algebraic. Let $A(z)$ be the maximum of $|A_j(z)|$, and let

$$\mu(r) = \frac{1}{2k\pi} \int_0^{2\pi} \log A(re^{i\varphi}) d\varphi.$$

Then $T(r, f) = \mu(r) + O(1)$. As the second fundamental theorem we have

$$\begin{aligned} \sum_{v=1}^q N(r, w_v) &> (q-2)T(r, f) - N(r, \mathfrak{Z}) \\ &\quad + \sum_{v=1}^q N_1(r, w_v) + O(\log r T) \\ &> (q-2k)T(r, f) \\ &\quad + \sum_{v=1}^q N_1(r, w_v) + O(\log r T), \end{aligned}$$

where $N_1(r, w)$ is the logarithmic integral of $n_1(r, w)$ which is the sum of the multiplicity minus one of all the roots of $f(x) - w = 0$ in \mathfrak{Z}_r . Furthermore, the **deficiency**, **ramification index** of $f(z)$, and **ramification index** of the surface \mathfrak{Z} are defined by

$$\delta(w) = 1 - \limsup N(r, w)/T(r, f)$$

$$= \liminf m(r, w)/T(r, f),$$

$$\mathfrak{D}(w) = \liminf N_1(r, w)/T(r, f),$$

$$\xi = \liminf N(r, \mathfrak{Z})/T(r, f).$$

With these notations, we have

$$\sum \delta(w_v) + \sum \mathfrak{D}(w_v) \leq 2 + \xi \leq 2k.$$

These results contain the Picard theorem and the Borel theorem. Furthermore, by considering the effect of branch points, the Ahlfors theory of covering surfaces can be extended to algebroidal functions (Y. Tsumura). By using this result, Bloch's theorem can be obtained very simply.

But these results do not contain the Varopoulos result. Taking into consideration the degeneracy index λ , H. Cartan proved the following inequality when $\lambda = 0$:

$$(q - k - 1)T(r, f) < \sum_{v=1}^q N_k(r, w_v) + S(r),$$

and for general λ he conjectured that the number $q - k - 1$ may be changed to the number $q - k - \lambda - 1$. This conjecture is still unsolved except for some special cases where for example, $q = k + \lambda + 2$ (N. Toda, *Nagoya Math. J.*, 91 (1983), 37–47).

Interesting results follow upon a study of this degeneracy index. For example, when f is an entire algebroidal function, if $\sum_{v=1}^{2k-1} \delta(w_v) > 2k - 2$ ($w_v \neq \infty$), there are at least $k - 1$ \dagger Picard's exceptional values in $\{w_v\}$ (Niino, Ozawa, and Toda).

With respect to the relation between the number of exceptional values and the order of fin $|z| < \infty$, there are some results similar to those for single-valued functions. For example, if f has $k + 1$ Picard's exceptional values, the order ρ must be a positive integer or ∞ .

D. Asymptotic Values and Other Results

In the single-valued case, Valiron, L. Ahlfors, W. H. J. Fuchs, A. Edrei, W. K. Hayman, and others studied the \dagger Borel direction, the number of \dagger asymptotic values, the relation between the deficiency values and the asymptotic values, etc. However, almost none of the corresponding results holds for the algebroidal case as shown by several counterexamples.

There is no relationship between the order of an entire algebroidal function and the number of its finite \dagger asymptotic values, which is quite different from the single-valued case. Furthermore, it is possible to have an infinite number of asymptotic values even if the order is equal to zero. If an algebroidal function f satisfies $\liminf T(r, f)/(\log r)^2 < +\infty$, then it has at most k asymptotic values (Valiron-Tumura). The Ahlfors theorem, which is concerned with the number of \dagger direct transcendental singular points of the inverse function and the order of a meromorphic function in the single-valued case, was extended to the algebroidal function by Lü Yinian (*Scientia Sinica*, 23 (1980)).

The \dagger Julia direction or the Borel direction for an algebroidal function is defined not on \mathbb{C} but on $|z| < cc$ because of the appearance of branch points. With respect to the Julia direction, there are results similar to those for the single-valued case, but it is unknown in general whether the Borel direction exists. A. Rauch proved that if $\int^\infty T(r, f)/r^{\rho+1} dr = \infty$,

18 B

Almost Periodic Functions

then there is a sector with an angle of at least π/ρ in which $L(\varphi) = \int^\infty \log^+ F(re^{i\varphi})/r^{\rho+1} dr$ diverges. Apart from the theory of distribution of values, Selberg obtained some conditions under which the inverse functions of \dagger Abelian integrals of a special kind reduce to algebroidal functions.

References

- [1] G. Rémoundos, Extensions aux fonctions algébroides multiformes du théorème de M. Picard et de ses généralisations, Gauthier-Villars, 1927.
- [2] H. Selberg, Algebroiden Funktionen und Umkehrfunktionen Abelscher Integrale, Avhandl. Norske Videnskaps-Akad., Oslo, 8 (1934), 1–72.
- [3] N. Toda, Surrounding algebroidal functions (in Japanese), Bull. Math. Soc. Japan, 24 (1972), 2006209.
- [4] H. Cartan, Sur les zéros des combinaisons linéaires de p fonctions holomorphes données, *Mathematica*, 7 (1933), 5–31.

18 (X.25)

Almost Periodic Functions

A. History

The theory of almost periodic functions was originated by H. Bohr in 1924 as a result of his study of \dagger Dirichlet series. The theory provides a method of studying a wide class of trigonometric series (\rightarrow 159 Fourier Series) of general type. Further generalizations were made by N. Wiener, V. V. Stepanov, A. S. Besikovich, S. Bochner, and others. H. Weyl, J. von Neumann, and others clarified the relations between this theory and \dagger representation of groups, specifically, the relations between almost periodic functions in a \dagger topological group and representation theory of a \dagger compact group.

B. Almost Periodic Functions in the Sense of Bohr

Let $f(x)$ be a complex-valued continuous function defined for all real values of x . A number τ is called a **translation number** of $f(x)$ belonging to $\varepsilon > 0$ if

$$\sup_{-\infty < x < \infty} |f(x + \tau) - f(x)| \leq \varepsilon.$$

If for any $\varepsilon > 0$ there exists a number $l(\varepsilon) > 0$

such that any interval of length $l(\varepsilon)$ contains a translation number off belonging to ε , then $f(x)$ is called **almost periodic in the sense of Bohr**. We denote by \mathbf{B} the set of all almost periodic functions in the sense of Bohr.

If $f(x)$ is \dagger periodic with a period p , then $f \in \mathbf{B}$, because each number $l \geq p$ plays the role of $l(\varepsilon)$ for any $\varepsilon > 0$. Any $f \in \mathbf{B}$ is bounded and uniformly continuous. A necessary and sufficient condition for a bounded continuous function on $(-\infty, \infty)$ to belong to \mathbf{B} is that for any given sequence $\{h_n\}$ of real numbers, there exist a subsequence $\{h_{n_k}\}$ such that the sequence of functions $\{f(x + h_{n_k})\}$ is uniformly convergent in $(-\infty, \infty)$; i.e., the set $\{f(x + h) \mid h \in (-\infty, \infty)\}$ is \dagger totally bounded with respect to the uniform norm $\|f\| = \sup |f(x)|$ in the space of bounded continuous functions in $(-\infty, \infty)$.

If $f(x) \in \mathbf{B}$, then $f(-x)$, $f(x)$, $\alpha f(x)$ (where α is a complex number), and $f(x + h)$ (where h is a real number) $\in \mathbf{B}$. If $f(x)$, $g(x) \in \mathbf{B}$, then $f(x) \pm g(x)$ and $f(x)g(x) \in \mathbf{B}$. If $f_n(x) \in \mathbf{B}$ and $\{f_n(x)\}$ converges uniformly to $f(x)$, then $f(x) \in \mathbf{B}$. For any real number A , $\exp i\lambda x$ (where i is the \dagger imaginary unit) is continuous and periodic. Hence the polynomial function $\sum_{n=1}^m \alpha_n \exp i\lambda_n x \in \mathbf{B}$. Moreover, if the latter function converges uniformly to $\sum \alpha_n \exp i\lambda_n x$ as m tends to ∞ , then the limit function is also an element of \mathbf{B} . The polynomial $\sum_{n=1}^m \alpha_n \exp i\lambda_n x$ and the series $\sum_{n=1}^{\infty} \alpha_n \exp i\lambda_n x$ are called a **generalized trigonometric polynomial** and a **generalized trigonometric series**, respectively.

For any $f \in \mathbf{B}$, its **mean** exists:

$$M[f] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_a^{a+T} f(x) dx.$$

The convergence of the right-hand formula is uniform in $a \in (-\infty, \infty)$, and the limit is independent of the choice of a . Thus $M[f]$ is a \dagger linear functional defined on \mathbf{B} . Since $M[\exp i\lambda x] = 1$ for $\lambda = 0$ and $= 0$ for $\lambda \neq 0$, the family $\{\exp i\lambda x \mid -\infty < \lambda < \infty\}$ is an \dagger orthonormal system with respect to the \dagger inner product $(f, g) = M[f(x)g(x)]$ defined on \mathbf{B} . Let $\alpha(\lambda) = M[f(x)\exp(-i\lambda x)]$ for any $f \in \mathbf{B}$; then there exist countably many values of λ for which $\alpha(\lambda)$ differs from zero. Denote these values of λ by $\lambda_1, \lambda_2, \dots$ and write $\alpha(\lambda_n) = \alpha_n$. We call the numbers $\alpha_1, \alpha_2, \dots, \alpha_n$, the **Fourier coefficients** of $f(x)$. The formal series $\sum_{n=1}^{\infty} \alpha_n \exp i\lambda_n x$ is called the **Fourier series** of $f(x)$. Moreover, the **Parseval equality** $M[|f(x)|^2] = \sum_{n=1}^{\infty} |\alpha_n|^2$ is valid for any $f \in \mathbf{B}$. For every periodic function, these definitions coincide with the ordinary Fourier coefficients and the Fourier series (\rightarrow 159 Fourier Series). Any almost periodic function in the sense of Bohr is uniquely determined by its Fourier

coefficients; i.e., if two almost periodic functions have the same Fourier series, then they are identical. For any $f \in \mathbf{B}$, its Fourier series does not always converge uniformly, but $f(x) \in \mathbf{B}$ can be approximated uniformly by a sequence of trigonometric polynomials. Hence the almost periodic functions in the sense of Bohr are also called **uniformly almost periodic functions**.

C. Generalizations of Almost Periodic Functions

Let $C(-\infty, \infty)$ be the space (- 168 Function Spaces) of all bounded continuous functions on $(-\infty, \infty)$ with distance $\rho(f, g) = \sup_{-\infty < x < \infty} |f(x) - g(x)|$. Then a uniformly almost periodic function is the limit of a sequence of trigonometric polynomials with respect to this distance. Generally, let ρ be a \dagger distance function introduced in a function space (whose elements are not necessarily continuous in $(-\infty, \infty)$). Then the limit of a sequence of generalized trigonometric polynomials with respect to the distance ρ is called an **almost periodic function with respect to ρ** . For example, for $p \geq 1$, we set

$$D_{Sp}[f, SI] = \sup_{-\infty < a < \infty} \left\{ \int_a^{a+1} |f(x) - g(x)|^p dx \right\}^{1/p},$$

$$D_{Wp}[f, SI] = \lim_{l \rightarrow \infty} \sup_{-\infty < a < \infty} \left\{ \frac{1}{l} \int_a^{a+l} |f(x) - g(x)|^p dx \right\}^{1/p}$$

These are distance functions. The properties of the corresponding almost periodic functions and their relations to other classes of almost periodic functions have been studied by Besikovich [1].

D. Analytic Almost Periodic Functions

Let D be a strip domain, $a < \operatorname{Re} z < b$, defined in a complex plane. For any \dagger holomorphic function $f(z)$ in D and $\varepsilon > 0$, a real number τ is called a **translation number** off belonging to $\varepsilon > 0$ if $\sup_{z \in D} |f(z + i\tau) - f(z)| \leq \varepsilon$. If for any $\varepsilon > 0$ there exists a number $l(\varepsilon) > 0$ such that any interval of length $l(\varepsilon)$ contains a translation number off belonging to ε , then $f(x)$ is called an **analytic almost periodic function** in D . We denote the set of all analytic almost periodic functions in $D = \{a < \operatorname{Re} z < b\}$ by $\mathbf{A}(a, b)$. If we fix an x in $a < x < b$, then $g(y) = f(x + iy)$ for any $f(z) \in \mathbf{A}(a, b)$ belongs to \mathbf{B} .

For any $f \in \mathbf{A}(a, b)$ there corresponds a \dagger Dirichlet series $\sum_{n=1}^{\infty} \alpha_n \exp \lambda_n z$ such that two analytic almost periodic functions are identical if the corresponding Dirichlet series are

identical. Here the coefficients

$$\alpha_n = M_y[f(x+iy)\exp(-iy, y)]$$

are determined independently of x ($a < x < b$), and Parseval's identity

$$M_y[|f(x+iy)|^2] = \sum_{n=1}^{\infty} |\alpha_n|^2 \exp 2\lambda_n x$$

holds (Bohr [3]). If the series

$$\sum \alpha_n \exp \lambda_n x \exp i\lambda_n y$$

at $x = a$ and $x = b$ represent the Fourier series of $f_a(y)$ and $f_b(y) \in \mathbf{B}$, respectively, then there exists $f \in \mathbf{A}(a, b)$ such that $f(z)$ is continuous on \bar{D} and

$$f(a+iy) = f_a(y), \quad f(b+iy) = f_b(y).$$

The behavior of $f \in \mathbf{A}(a, b)$ at the boundary or exterior points of the domain $D = \{a < \operatorname{Re} z < b\}$ has also been investigated by Besikovich [1].

E. Almost Periodic Functions on Groups

Von Neumann defined almost periodic functions on any group, generalizing the characterization of uniformly almost periodic functions on $(-\infty, \infty)$. Let $B(G)$ be the set of all complex-valued bounded functions on a group G . Then $B(G)$ is a metric space with the distance $\rho(f, g) = \sup_{x \in G} |f(x) - g(x)|$. If for any $f \in B(G)$ the set $A_f = \{f_{a,b}(x) = f(axb) | a, b \in G\}$ is totally bounded in the metric space $B(G)$, we call f an **almost periodic function on the group** G . This condition is equivalent to the total boundedness of $B_f = \{f_a(x) = f(xa) | a \in G\}$ or $C_f = \{f_a(x) = f(ax) | a \in G\}$. We denote the set of almost periodic functions on G by $\mathcal{A}(G)$.

For $f(x), g(x) \in \mathcal{A}(G)$, the linear combinations $a \cdot f(x) + b \cdot g(x)$ ($a, b \in \mathbb{C}$) and the product $f(x)g(x)$ are both contained in $\mathcal{A}(G)$. If $f_n \in \mathcal{A}(G)$ and $\{f_n\}$ converges to f uniformly on G , then $f \in \mathcal{A}(G)$. If $f \in \mathcal{A}(G)$, then $f_{a,b}, f_{a,b}f \in \mathcal{A}(G)$ also. Hence $\mathcal{A}(G)$ is a closed subalgebra invariant under two-sided translation in the † Banach algebra $l^{\infty}(G)$. For any $f \in \mathcal{A}(G)$ there exists only one number $M[f]$ in the closure (with respect to the distance ρ in $B(G)$) of

$$A'_f = \left\{ \sum_{i=1}^n c_i f(a_i x b_i) \mid c_i > 0, \sum c_i = 1, a_i, b_i \in G \right\}$$

(= the least closed † convex set including A_f). We call $M[f]$ the **mean** off on G . The mapping $f \rightarrow M[f]$ is a linear functional on $\mathcal{A}(G)$, and we have $M[f] \geq 0$ if $f \geq 0$.

F. Relation to Bounded Representation

Suppose that we are given a finite-dimensional matrix representation $\mathbf{D}(x) = (d_{ij}(x))$ of a group

G . Then the following three conditions are equivalent: (i) All the $d_{ij}(x)$ are bounded on G . (ii) All the $d_{ij}(x)$ are almost periodic in G . (iii) The representation D is † equivalent to a representation by unitary matrices. The inner product $(f, g) = M[f(x)g(x)]$ provides the algebra $\mathcal{A}(G)$ with the structure of a † pre-Hilbert space. Let $H(G)$ be the Hilbert space that is the completion of $\mathcal{A}(G)$. If we select $D''(x) = (d''_{ij}(x))$ from each L , where L is an equivalence class of bounded irreducible representations of G , and if n_{λ} is the order of D^{λ} , then $\{(1/\sqrt{n_{\lambda}})d''_{ij}(x) | 1 \leq i, j \leq n_{\lambda}, \lambda \in L\}$ is a † complete orthonormal system in the Hilbert space $H(G)$. Any $f(x) \in \mathcal{A}(G)$ can be approximated uniformly in G by a finite linear combination of the $d_{ij}(x)$.

G. Almost Periodic Functions on Topological Groups

When G is a † separated topological group, we denote the set of all continuous functions on G contained in $\mathcal{A}(G)$ by $\mathcal{A}_{*}(G)$. The statements of the theorems in the previous section concerning $\mathcal{A}(G)$ and the representation D remain valid if we replace $\mathcal{A}(G)$ by $\mathcal{A}_{*}(G)$ and replace D by a continuous representation of G . In particular, if G is the additive group of real numbers \mathbf{R} , then $\mathbf{d}(\mathbf{R})$ is exactly \mathbf{B} .

H. Relation to Compact Groups

Every continuous function on a compact group G is almost periodic; i.e., $\mathcal{A}_{*}(G) = C(G)$. The mean value $M[f]$ of $f \in \mathcal{A}_{*}(G)$ is identical to $\int_G f(x) dx$, where the † Haar measure dx is normalized so that $\int_G dx = 1$. In this case, the theory of bounded representations discussed above is the Peter-Weyl theory (- 69 Compact Groups).

In general, let G be a separated topological group. There exists a continuous homomorphism φ of G onto a compact group $K = K(G)$ with the following two properties: (i) For any compact group K' and a continuous homomorphism $\varphi': G \rightarrow K'$, there exists a continuous homomorphism $\psi: K \rightarrow K'$ such that $\varphi' = \psi \circ \varphi$. (ii) Such a pair $\mathbf{K} = (K, \varphi)$ is unique up to isomorphism. K is called the **Bohr compactification** of G , and φ is called the canonical mapping. In particular, suppose that G is a locally compact Abelian group and G^* is its † character group. We denote by G' the group G^* with discrete topology. Let K be the character group of G' , and let φ^* be the identity mapping $G' \rightarrow G^*$ and φ be its conjugate mapping $G \rightarrow K$, which is a continuous homomorphism. Then K is the Bohr compactification of G with the canonical mapping φ . A necessary and sufficient condition for f on G to be con-

tinuous almost periodic is that a continuous function f' on K such that $f' = f' \circ \varphi$ must exist. If this condition is satisfied, then the mean $M[f]$ is identical to $\int_K f'(x) dx$. For any finite-dimensional continuous unitary representation D' of K , $D = D' \circ \varphi$ is a finite-dimensional continuous unitary representation of G , and vice versa. Hence there exists a canonical isomorphism (determined by $D = D' \circ \varphi$) between the equivalence classes of finite-dimensional unitary representations of a separated topological group G and the equivalence classes of finite-dimensional unitary representations of its Bohr compactification K . The kernel of the canonical mapping $\varphi: G \rightarrow K$ is identical to the intersection of all the kernels of finite-dimensional continuous unitary representations of G .

1. Maximally Almost Periodic Groups

Let G be a topological group. If for each pair a, b of distinct elements of G there exists a continuous almost periodic function f on G such that $f(a) \neq f(b)$, then G is called a **maximally almost periodic group**. This is the case if and only if G has sufficiently many finite-dimensional unitary representations. For a connected locally compact group G , the following six conditions are equivalent: (1) G is a maximally almost periodic group. (2) There is a one-to-one continuous homomorphism from G into a compact group. (3) G is the direct product of a compact group and a vector group \mathbf{R}^n . (4) G is the projective limit of Lie groups that are locally isomorphic to compact groups. (5) The quotient group G/Z is compact, where Z is the center of G . (6) The system of all neighborhoods that are invariant under the inner automorphisms constitutes a basis for the neighborhood system of the unit [7].

Moreover, any discrete free group is maximally almost periodic. If there is no continuous almost periodic function except constant functions, the topological group is called **minimally almost periodic**. Any noncompact connected simple Lie group is minimally almost periodic.

References

- [1] A. S. Besicovich (Besikovich), *Almost periodic functions*, Cambridge Univ. Press, 1932.
- [2] S. Bochner, *Bertrager zur Theorie der fast-periodischen Funktionen I, II*, Math. Ann., 96 (1926), 119–147, 383–409.
- [3] H. Bohr, *Fastperiodischen Funktionen*, Erg. Math., Springer, 1932.

[4] V. Stepanov, *über einige Verallgemeinerungen der fastperiodischen Funktionen*, Math. Ann., 95 (1926) 473–498.

[5] N. Wiener, *On the representation of functions by trigonometrical integrals*, Math. Z., 24 (1926), 575–616.

[6] J. von Neumann, *Almost periodic functions in a group I*, Trans. Amer. Math. Soc., 36 (1934), 445–492.

[7] J. Dixmier, *Les C*-algèbres et leurs représentations*, Gauthier-Villars, 1964.

[8] W. Maak, *Fastperiodische Funktionen*, Springer, 1950.

19 (XV.1 0) Analog Computation

A. History

The term **analog computation** is a generic term describing various techniques of computation employing diagrams, or physical systems whose equations are similar to the mathematical problems in question. The history of analog computation is probably as old as that of digital computation; for example, the ancient Greeks tried to solve cubic equations using diagrams, and the astrolabe widely used by astronomers through the medieval period is also a kind of analog computer. Soon after the discovery of logarithms, the **slide rule** was invented. In the 18th Century, the **planimeter**, used to measure plane areas, was introduced, and in the 19th Century, the **nomogram** appeared (- Section D). In the first half of 20th Century, a large electronic analog computer was developed, thus predating the first practical digital computer.

However, analog computation has an essential defect, namely, the limitation of accuracy. Today it is useful for simple calculation, but is rapidly becoming less important as the development of digital computers, including pocket calculators, advances.

B. Graphical Calculation

Graphical calculation is a method of computation by means of geometric constructions using common drawing tools. Some typical examples of practical graphical calculation are the following: evaluation of linear functions of several variables (J. Massau, 1887), of systems of linear equations (F. J. van den Berg, 1888), of polynomials (J. A. Segner, 1761), of algebraic equations (Lill, 1867), **graphical integration**, **graphical differentiation**, and the solu-

tion of ordinary differential equations of the first order (J. Massau, 1878), of linear differential equations (Czuber), and of ordinary differential equations of the second order (Lord Kelvin, 1892).

Using bisquare-root graph paper, sold under the names of **binomial probability paper** or **stochastic paper**; we can handle F -distributions or other probability distributions reducible to F -distributions, such as binomial or normal distributions, to a fairly good degree of approximation. This **graphical method of statistical inference** is, even now, a powerful method of statistical quality control.

C. Graphical Mechanics

Graphical mechanics is the graphical treatment of mechanical problems, especially problems of equilibrium. In this method, the fundamental constructions are the composition and resolution of forces by means of force polygons (Fig. 1). This method is also applicable to problems in dynamics when they can be reduced to problems of equilibrium by means of d'Alembert's principle.

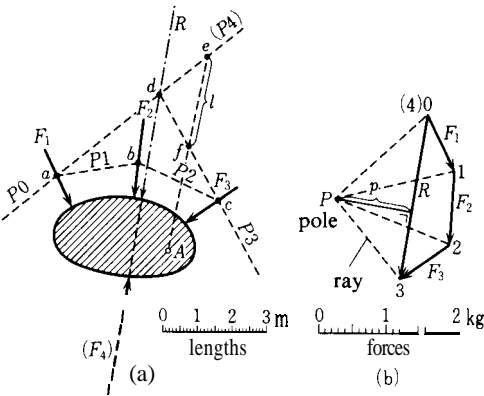


Fig. 1

An example of the graphical method of constructing the composite and the line of action of three given forces F_1 , F_2 , and F_3 . In the force polygon in (b), the vector $\vec{03}$ represents the composite R . On the link polygon in (a), the point d is a point on the line of action of the force. If the fourth force F_4 has the same magnitude as R with opposite direction, both polygons are closed, and the four forces F_1 , F_2 , F_3 , and F_4 are in equilibrium.

Graphical mechanics is most convenient when applied to problems reducible to 2-dimensional structural mechanics. But we can also use it for 3-dimensional problems if we work on projections, e.g., on the plan and the elevation, of the original body. In recent years, however, the amount of work involved in geometric construction has been deemed non-negligible, and the technique of computer

graphics is now considered to be much more convenient. Here the computation is done by a digital computer, and the result is displayed in a convenient graphical form.

D. Nomograms

Nomograms are charts in which we can easily read off the corresponding value u_n from given values u_1, \dots, u_{n-1} when there is a relation $F(u_1, \dots, u_n) = 0$ among n real variables u_1, \dots, u_n . The construction of nomograms has been thoroughly investigated by M. d'Ocagne [5].

A function of two variables $F(u_1, u_2) = 0$ can be represented by a **two-sided scale** or by **functional paper**. The most useful nomograms are those for functions of three variables, $F(u_1, u_2, u_3) = 0$. Changing this into $f_i(x, y, u_i) = 0$ ($i = 1, 2, 3$) and drawing the curves $f_i(x, y, u_i) = 0$ for fixed values of u_i , we have an **intersection chart**, where three curves for the corresponding values u_1, u_2, u_3 meet at a point. In practice, however, it is much more convenient to use the **alignment chart**, which is a dual of the intersection chart. It is especially useful when the relation $F(u_1, u_2, u_3) = 0$ can be decomposed into an equation of the form

$$\begin{vmatrix} f_1(u_1) & g_1(u_1) & h_1(u_1) \\ f_2(u_2) & g_2(u_2) & h_2(u_2) \\ f_3(u_3) & g_3(u_3) & h_3(u_3) \end{vmatrix} = 0. \quad (1)$$

Putting $x_i = f_i/h_i$, $y_i = g_i/h_i$ ($i = 1, 2, 3$) and drawing the curves (x_i, y_i) with the parameter u_i scaled on it (u_i -scaled), the corresponding values u_1, u_2 , and u_3 satisfying relation (1) lie on a straight line (Fig. 2). Using this property we can easily read off the corresponding values (by laying down a ruler). In the strict sense, the term "nomogram" usually refers specifically to an alignment chart.

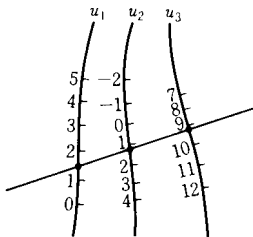


Fig. 2

For four or more variables, we can apply similar techniques if the function is **separable** into functions of three variables. If it is not separable, we can sometimes apply such techniques such as two-functional scales with functional networks, cocircular charts, coplanar charts, or moving charts. Using func-

tional approximations, an approximate nomogram has recently been constructed with errors considered to be admissible in practical use.

Historically, the origin of Hilbert's 13th problem, which asks if it is possible to represent a function of many variables as the composite of functions of two variables, comes from a study of nomograms with many variables.

E. Analog Computers

The **analog computer** is a special purpose machine designed for a specified analog computation. In a wider sense it includes special devices that perform [†]analog simulation. The following mechanical devices are well known: the **pantograph** for copying plans, the **harmonic analyzer** for obtaining the Fourier expansion of a periodic function, and V. Bush's **differential analyzer**. Since about 1940 large **electronic analog computers** for solving differential equations have been extensively developed.

Analog computers have up to recently had the advantages of ease of construction, simplicity, and inexpensive operation, and they were also considered to be fast enough for use in **real-time** computation. To compensate for the limitations of analog computers in comparison to digital ones, several **hybrid computers** have also been used. However, with the rapid progress of digital computers and digital-analog converters, the analog computers are now considered less important than in the past.

F. Curve Fitting

Curve fitting is a method of finding a simple curve $y = f(x)$ supplying the best possible approximation to the values y_1, y_2, \dots, y_n for discrete values x_1, x_2, \dots, x_n of the independent variable. A polynomial passing through all the given points is constructed by means of [†]tinterpolation. For experimental data we usually construct a curve by using methods, such as the [†]method of least squares, that take the errors due to observation into account.

A function constructed to best fit the observed values of the function $y = f(x)$ expressing some physical law is called an **empirical formula**, in contrast to the **theoretical formula**. Usually, we first assume that the function contains several **empirical constants**, which are then determined to fit the experimental data.

The most common case is a linear approximation (using a linear function) with a suitable change of variables. **Semilogarithmic**

paper, **logarithmic paper**, or **probability paper** are used to facilitate transformations in terms of logarithms or normal distributions.

We also frequently use polynomials or trigonometric polynomials of lower degree. To determine the empirical constants the following methods are used: the graphing method, the selection of certain points, the mean value method, or the method of least squares.

G. Orthogonal Polynomials

Let there be given the values y_0, y_1, \dots, y_{n-1} for discrete values $x = 0, 1, \dots, n-1$ with equal differences. A polynomial $f(x)$ of degree k ($\leq n$) that minimizes the sum of the squares $G_k(n) = \sum (y_m - f(m))^2$ is given by

$$f(x) = \sum_{v=0}^k a_v q_v(n, x),$$

$$a_v = \sum_{m=0}^{n-1} y_m q_v(n, m) / S_v(n),$$

$$S_v(n) = \sum_{m=0}^{n-1} (q_v(n, m))^2,$$

and we have $G_k(n) = \sum_{m=0}^{n-1} y_m^2 - \sum_{v=0}^k a_v^2 S_v(n)$. Here the function $q_v(n, x)$, called a **Chebyshev q-function**, is defined by

$$q_v(n, x) = \sum_{m=0}^v \binom{v+m}{m} \binom{v-n}{v-m} \binom{x}{m}$$

In practical applications, it is better to replace $q_v(n, x)$ by the function

$$q_v^*(n, x) = q_v(n, x) / 2^{-v} v! M_v(n),$$

where $M_v(n)$ is the greatest common divisor of

$$\binom{v+m}{m} \binom{n-1-m}{v-m} \quad (m=0, 1, \dots, v).$$

The values $q_v^*(n, m)$ ($m=0, 1, \dots, n-1$) are mutually coprime integers. The function $q_v^*(n, m)$ is called the **simplest Chebyshev q-function** or **simplest orthogonal polynomial**, and is sometimes denoted by $X_{n,v}(x)$, $\xi'_{n,m}(x)$, or $\varphi_{n,m}(x)$.

When two functions satisfy the condition

$$(f, g)_n = \sum_{m=0}^n f(m)g(m) = 0,$$

they are called **orthogonal for a finite sum**. If from $1, x, x^2, \dots, x^v$ we construct a system orthogonal with respect to this definition, then we have the polynomial

$$P_{v,n}(x) = \sum_{m=0}^v (-1)^m \binom{v}{m} \binom{v+m}{m} \binom{x}{m} / \binom{n}{m},$$

which has the following connection with q_v :

$$q_v(n, x) = ((-1)^v (n-1)! / 2^v (n-v-1)!) P_{v,n-1}(x).$$

The polynomial $P_{v,n}(x)$ is called a **Chebyshev orthogonal polynomial** or sometimes simply an **orthogonal polynomial**. The polynomials $P_{v,n}(x)$ are orthogonal for the finite sum. If $n \rightarrow \infty$ and $0 \leq x \leq 1$, the function $P_{v,n}(x)$ tends to $P_v(1 - 2x)$, where P_v is the \dagger Legendre polynomial. These functions may be conveniently used in least-squares curve fitting.

References

- [1] G. W. Smith and R. C. Wood, *Principles of analog computation*, McGraw-Hill, 1959.
For graphical calculation,
[2] C. Runge and F. A. Willers, *Numberische und graphische Quadratur und Integration gewöhnlicher und partieller Differentialgleichungen*, Enzykl. der Math., Bd. II, Teil 3, Heft 2 (1915), 477176.
For graphical mechanics,
[3] C. B. Biezeno and R. Grammel, *Technische Dynamik*, Springer, 1939.
[4] H. F. B. Müller-Breslau, *Die graphische Statik der Baukonstruktionen*, Kröner, I, sixth edition, 1927; II, pt. 1, fifth edition, 1922; II, pt. 2, second edition, 1925.
For nomograms,
[5] M. d'Ocagne, *Traité de nomographie*, Gauthier-Villars, second edition, 1921.
[6] R. Soreau, *Nomographie*, E. Chiron, second edition, 1921.
[7] H. Schwerdt, *Lehrbuch der Nomographie*, Springer, 1924.
[8] D. P. Adams, *An index of nomograms*, MIT Press, 1950.
For analog computers,
[9] G. A. Korn and T. M. Korn, *Electronic analog computers*, McGraw-Hill, 1952; second edition, 1956.
For curve fitting and empirical formulas,
[10] T. R. Running, *Empirical formulas*, Wiley, 1917.
[11] P. G. Guest, *Numerical methods of curve fitting*, Cambridge Univ. Press, 1961.
For tables of Chebyshev q functions,
[12] R. A. Fisher and F. Yates, *Statistical tables for biological, agricultural, and medical research*, Oliver & Boyd, 1938 [$n \leq 75, v \leq 5$].
[13] R. L. Anderson and E. E. Houseman, *Tables of orthogonal polynomial values extended to $n = 104$* , Agricultural Exp. Station, Iowa State College Stat. Sec. Res. Bull., 297 (1942), 595-672 [$n \leq 104, v \leq 5$].
[14] J. Yamanouchi, *On the Chebyshev q functions. I. Simplest Chebyshev q functions and their tables* (in Japanese), Report of the Electrotechnical Laboratory, 454 (1942) [$n \leq 15, v \leq n-1; 16 \leq n \leq 21, v \leq 9; 22 \leq n \leq 31, v \leq 7; 32 \leq n \leq 51, v \leq 5$].

20 (X.1) Analysis

The origin of analysis can be traced back to the time when Eudoxus (4th century B.C.) and Archimedes (3rd century B.C.) devised the so-called method of exhaustion for calculating the area of a plane figure and the volume of a solid. Their objects of investigation were restricted, however, to particular types of figures or solids. In the 16th and 17th centuries, F. Viète, J. Kepler, and B. Cavalieri again took up this problem. In the 17th Century, the problem of drawing a tangent to a given curve was studied by R. Descartes, P. de Fermat, B. Pascal, and J. Wallis. Fermat, in particular, applied the result to find the maxima and minima of certain functions. It is worth noting that a certain type of mathematics powerful enough to produce similar results was independently developed in Japan around that time. In 1684, G. Leibniz in Germany introduced the symbols dx and dy in treating the same problem. He proved that dy/dx represents the slope of the tangent to the curve at a given point and discovered a new operation to calculate it. In 1686, he established the "inverse tangent method," which is what we now call integral calculus. He also introduced the notation \int . On the other hand, I. Newton in England developed his "method of fluxions," corresponding to our differential and integral calculus, from the viewpoint of mechanics. But neither Leibniz nor Newton formulated the fundamental concepts rigorously, and therefore they were criticized severely by many contemporary scholars. The new calculus gained ground in Great Britain rather slowly. B. Taylor in England and C. MacLaurin in Scotland demonstrated its usefulness in 1715 and in 1745, respectively. On the continent, however, Leibniz's symbolic calculus was taken up by mathematicians of the Bernoulli family, G. F. A. de l'Hôpital, G. Fagnano, and many others; with it, they solved many scientific problems which until then had remained intractable. This motivated subsequent researchers to pose new problems in the form of differential equations.

One of these problems, treated by J. le Rond d'Alembert in relation to the vibration of a chord, concerns the \dagger partial differential equation

$$\partial^2 y / \partial t^2 = a^2 \partial^2 y / \partial x^2 \quad (1)$$

for $y = y(t, x)$ with the boundary conditions $y = 0$ for $x = 0$ and $x = l$. He obtained the solution $y = f(at + x) - f(at - x)$, where f is an arbitrary function of period $2l$. In 1753, D.

Bernoulli showed that solutions of equation (1) are given by functions of the form

$$y = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos \frac{k\pi x}{l} + b_k \sin \frac{k\pi x}{l} \right)$$

These two kinds of solutions gave rise to the question of whether an arbitrary function can be expressed by a [†]trigonometric series. This problem was studied by A. C. Clairaut, J. L. Lagrange, and L. Euler. In 1807, J. Fourier in France, in treating a problem on the conduction of heat, claimed that an arbitrary function of period 2π can be expressed as

$$y = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (2)$$

where the coefficients a_k and b_k are given by

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx, \\ b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \, dx. \quad (3)$$

This series is now called the [†]Fourier series, but Fourier never verified the fact that the series (2) with coefficients (3) converges and represents $f(x)$. It was only as late as 1820 that A. L. Cauchy in France first noted that to treat a series properly, one must examine its convergence.

In the 19th Century, the concept of [†]functions, which had been taken in the sense of “analytic expressions,” came to be defined by the correspondence relation. Cauchy clarified the ideas of [†]limit and [†]continuity, [†]differentiability and [†]integrability. He showed that a function that is continuous in a bounded closed interval is integrable in that interval. But his proof was not rigorous, as he lacked the notion of [†]uniform continuity. In his 1854 paper on the trigonometric series, B. Riemann in Germany considered the integrability of functions that might be discontinuous and introduced the concept of what we now call the [†]Riemann integral.

The theory of [†]sets, initiated by G. Cantor in Germany in his paper of 1874, revolutionized analysis. R. Baire, E. Borel, and H. Lebesgue in France contributed to the establishment of analysis on the basis of set theory. Baire made a classification of discontinuous functions. Generalizing his results, Lebesgue gave a definition of analytic expressions, thus clarifying the term that had been used vaguely since the time of Euler. Lebesgue also tried to define the concepts of the integral of a function, the length of a curve, and the area of a surface from the most general viewpoint. In generalizing the notion of [†]measure introduced by Borel, he established in 1902 the theory of [†]Lebesgue measure with which he laid the

foundations of the theory of [†]Lebesgue integrals. The introduction of this theory gave to the theory of Fourier series a new turn in the direction of functional analysis. Measure theory was also employed by A. N. Kolmogorov of the Soviet Union to lay a solid foundation for [†]probability theory in 1933.

The study of functions of a complex variable was originated by Cauchy in the first half of the 19th Century. He began his research by introducing the notion of “monogenic functions”; a function which is monogenic at every point of a domain is what we now call a [†]holomorphic function. He established [†]Cauchy’s integral theorem and integral formulas for these functions, and deduced from these theorems the [†]residue theorem for functions with [†]poles. Making use of the integral formula, Cauchy proved that a function that is holomorphic at a point a can be expanded in a power series of the form $\sum_{k=0}^{\infty} a_k(z-a)^k$ in a neighborhood of this point.

Riemann considered a complex variable w as a function of another complex variable z when dw/dz is independent of the value of the differential dz . This amounts to the same thing as a “monogenic function” of Cauchy. Riemann’s mapping theorem became a model for subsequent developments. Riemann introduced the concept of [†]Riemann surfaces in order to [†]uniformize multivalued functions. This important idea was basic to the progress of analysis and [†]topology in the 20th Century.

K. Weierstrass, who was a contemporary of Riemann, developed the theory of functions of a complex variable from a purely analytic viewpoint. He defined an element of a function to be a power series $\sum_{k=0}^{\infty} a_k(z-a)^k$ of $z-a$, representing a holomorphic function in the interior of its [†]circle of convergence, and defined an analytic function to be an aggregate of such elements that are derived from one of them by means of [†]analytic continuations, along all curves having the point a as the initial point.

Riemann and Weierstrass constructed their theory to complete the theory of [†]elliptic functions and [†]Abelian functions that was initiated by N. H. Abel and C. G. J. Jacobi. Their results were a high point of 19th-century classical mathematics. H. Poincaré in France built upon their work. Another high point in the theory of functions of a complex variable was reached when J. Hadamard and C. de La Vallée Poussin in France made use of it to prove the [†]prime number theorem in 1896.

Weierstrass also initiated the study of functions of several complex variables and was succeeded by Poincaré, P. Cousin, and E. Picard in France. They tried to extend the theory of functions of one complex variable to

that of many variables. F. Hartogs, however, discovered a phenomenon quite different from the case of one variable. Cousin posed in 1895 the problem of constructing a function of several complex variables that has assigned poles. This problem was pursued by A. Weil in France and solved by K. Oka in Japan in 1936. Also, the problem of characterizing domains of holomorphy was investigated by E. E. Levi in Italy, H. Cartan in France, and P. Thullen and K. Stein in Germany and was finally solved by Oka in 1953. H. Cartan and J.-P. Serre reformulated these results in terms of cohomology with coefficients in sheaves. This considerably influenced the formulations of mathematics thereafter. These results were further generalized to analytic space by H. Grauert, R. Remmert, and Stein in Germany. The theory of complex manifolds, which are generalizations of Riemann surfaces to several variables, was initiated by W. V. D. Hodge in England and K. Kodaira in Japan and continued by F. Hirzebruch in Germany, M. F. Atiyah in England, and I. M. Singer in the United States.

Differential calculus gives a general method of finding extreme values of a given function. Likewise, in order to find a function that produces an extremal of the given functional, the calculus of variations was created in the 18th Century. For example, Euler considered the problem of finding a particular function $y(x)$ that renders the functional $\int_a^b F(x, y, y') dx$ an extreme value among all those functions $y(x)$ for which the plane curve $y = y(x)$ passes through two given points (a, A) and (b, B) of the plane; he showed that this $y(x)$ must satisfy the differential equation $dF_y/dx = F_y = 0$ (1744). Lagrange, W. R. Hamilton, and others developed this result into a general variational principle that governs not only classical mechanics but also quantum mechanics.

From research on the continuity or differentiability of the functional with respect to y emerged the idea of considering a function as a "point" in a function space. This gave rise to functional analysis, a branch of analysis that treats functions as elements of certain spaces and utilizes the methods of algebra and topology. The first result in this regard was the theory of integral equations of V. Volterra in Italy and E. I. Fredholm in Sweden in the beginning of the 20th Century. Fredholm's work was motivated by his desire to solve the Dirichlet problem, the solution of which had been used by Riemann in the proof of his mapping theorem, etc. However, Riemann's own proof of existence of the solution, called the Dirichlet principle, was not rigorous, and attempts to save the proof provided one of the central problems of analysis for some time.

Fredholm's solutions were different from Riemann's but D. Hilbert in Germany was able to justify Riemann's original proof. Hilbert's proof was later simplified and generalized by R. Courant, H. Weyl, and others. Hilbert also introduced the function spaces l_2 and L_2 to study the eigenvalue problem of integral equations with symmetric kernels. Later J. von Neumann in Germany established spectral theory in abstract Hilbert spaces and applied it to the mathematical foundations of quantum mechanics (1929). S. Banach in Poland created the theory of linear operators in Banach spaces (1932). This theory was further generalized to that of linear topological spaces and was applied to the theory of distributions.

In their study of partial differential equations and Fourier analysis, Hadamard, J. Leray, S. L. Sobolev, T. Carleman, and many others had to extend the notion of functions; they also enlarged the notion of derivatives. L. Schwartz in France introduced distributions and defined derivatives in the sense of distributions to unify these generalizations (1945). M. Sato in Japan defined more general generalized functions, called hyperfunctions (1958). It has become evident that both distributions and hyperfunctions have provided the most powerful tools in recent research in the general theory of partial differential equations, to which L. Hormander in Sweden has made outstanding contributions.

References

- [1] M. B. Cantor, *Vorlesungen über Geschichte der Mathematik*, Teubner, 1894/1908. Also \rightarrow references to 265 Mathematics in the 17th Century, 266 Mathematics in the 18th Century, and 267 Mathematics in the 19th Century.
- [2] N. Bourbaki, *Eléments de mathématique*, Fonctions d'une variable réelle, Integration, Espace vectoriels topologiques, Théorie spectrales, *Actualités Sci. Ind.*, Hermann, 1955–1969.
- [3] C. Carathéodory, *Vorlesungen über reelle Funktionen*, second edition, Teubner, 1927 (Chelsea, 1968).
- [4] H. Cartan, 1. Calcul différentiel, 2. Formes différentielles, Hermann, 1967.
- [5] G. Choquet, *Lectures on analysis I, II, III*, Benjamin, 1969.
- [6] R. Courant and D. Hilbert, *Methods of mathematical physics I, II*, Interscience, 1953, 1962.
- [7] R. Courant, *Differential- und Integralrechnung I, II*, Springer, third edition, 1955; English translation, *Differential and integral calculus*, Nordemann, 1938.

- [S] C. J. de La Vallée-Poussin, Cours d'analyse infinitésimale I, II, Gauthier-Villars, seventh edition, 1930.
- [9] J. Dieudonné, Treatise on analysis I–VII, Academic Press, 1960–1978.
- [10] J. Dieudonné, Eléments d'analyse I–VIII, Hermann, 1971–1978.
- [11] E. Goursat, Cours d'analyse mathématique I, II, III, Gauthier-Villars, fifth edition, 1933–1956.
- [12] J. Hadamard, Cours d'analyse I, II, Hermann, 1927, 1930.
- [13] G. H. Hardy, A course of pure mathematics, Cambridge Univ. Press, seventh edition, 1938.
- [14] D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Teubner, second edition, 1924.
- [15] C. Jordan, Cours d'analyse de l'école polytechnique: I. Calcul différentiel; II. Calcul intégral; III. Calcul intégral (équations différentielles), Gauthier-Villars, third edition, I, 1909; II, 1913; III, 1915.
- [16] H. L. Lebesgue, Leçons sur l'intégration et la recherche des fonctions primitives, Gauthier-Villars, second edition, 1928.
- [17] É. Picard, Traité d'analyse, Gauthier-Villars, I, 1922; II, 1926; III, 1928.
- [18] G. Pólya and G. Szegő, Aufgaben und Lehrsätze aus der Analysis I, 11, Springer, second edition, 1954.
- [19] W. Rudin, Real and complex analysis, McGraw-Hill, 1966.
- [20] L. Schwartz, Cours d'analyse mathématique I, II, Hermann, 1967.
- [20A] L. Schwartz, Analyse-Topologie générale et analyse fonctionnelle, Hermann, 1971.
- [21] L. Schwartz, Théorie des distributions, Actualités Sci. Ind., Hermann, 1950, second edition, 1966.
- [22] V. I. Smirnov, A course of higher mathematics: I. Elementary calculus; II. Advanced calculus; III-1. Linear algebra; III-2. Complex variables, special functions; IV. Integral equations and partial differential equations; V. Integration and functional analysis, Addison-Wesley, 1964. (Original in Russian, 1961.)
- Also → references to 106 Differential Calculus, 107 Differential Equations, 162 Functional Analysis, and 198 Holomorphic Functions.

21 (X1.20)

Analytic Functions of Several Complex Variables

A. Holomorphic Functions

As in the case of holomorphic functions of one complex variable, the definition of holo-

morphic functions can be given in two ways: the first definition utilizes differentiability, following the approach of B. Riemann; and the second method utilizes the notion of power series expansion as developed by K. Weierstrass. In this article, $N = \{0, 1, 2, \dots\}$.

B. Power Series

Let z be an n -tuple of complex variables z_1, \dots, z_n , and $c = (c_1, \dots, c_n)$ a point of \mathbb{C}^n . An infinite sum P of monomials $a_k(z - c)^k = a_{k_1, \dots, k_n}(z_1 - c_1)^{k_1} \dots (z_n - c_n)^{k_n}$ ($k = (k_1, \dots, k_n) \in \mathbb{N}^n$), where $a_k \in \mathbb{C}$, is called a **power series** with center c and coefficients a_k . If, for a bijection φ of \mathbb{N} onto \mathbb{N}^n , the simple series $\sum_{p \in \mathbb{N}} |a_{\varphi(p)}| (z - c)^{\varphi(p)}$ is convergent at $z = z^0$, we say that P is **absolutely convergent** at z^0 . Its sum at z^0 , denoted by $\sum a_k(z^0 - c)^k$, is defined as the sum $\sum a_{\varphi(p)}(z^0 - c)^{\varphi(p)}$, which is independent of the choice of φ . If P is uniformly bounded at z^0 , then P is absolutely convergent at every point of the open **polydisk** $S = \{z \mid |z_j - c_j| < |z_j^0 - c_j|, j = 1, \dots, n\}$. Furthermore, in this case P converges absolutely and uniformly on every compact set in S (N. H. Abel).

The **convergence domain** of a power series P is the set D of points z^0 such that P is absolutely convergent at every point in a neighborhood of z^0 . The interior of the set B of points at which the infinite sum P is uniformly bounded is equal to D . A **(complete) Reinhardt domain** with center c is a domain D in \mathbb{C}^n such that whenever D contains z_0 , the domain D also contains the torus $\{z \mid |z_j - c_j| = |z_j^0 - c_j|, j = 1, \dots, n\}$ (the closed polydisk $\{z \mid |z_j - c_j| \leq |z_j^0 - c_j|, j = 1, \dots, n\}$). If the convergence domain D of the power series P is not empty, it is a complete Reinhardt domain and is also **logarithmically convex**; that is, the set $D - \bigcup_j \{z \mid z_j = c_j\}$ is mapped onto a convex domain in \mathbb{R}^n by the mapping $z_j \rightarrow \log |z_j - c_j|$ ($j = 1, \dots, n$). The set \tilde{D} of points at which P is absolutely convergent is, in general, greater than D , and it is possible that \tilde{D} contains exterior points of D . A **thorn** of D is the set of exterior points of D contained in \tilde{D} that are located on the planes $\{z \mid z_j = c_j\}$ ($j = 1, \dots, n$). An n -tuple $r \in \mathbb{R}_+^n$ is called a set of **associated convergence radii** if P is absolutely convergent at every point of $\{z \mid |z_j - c_j| < r_j, j = 1, \dots, n\}$ but not of $\{z \mid |z_j - c_j| > r_j, j = 1, \dots, n\}$. An n -tuple of associated convergence radii may not be uniquely determined, but it satisfies

$$\limsup_{|k| \rightarrow +\infty} (|a_k| r^k)^{1/|k|} = 1, \quad |k| = k_1 + \dots + k_n$$

(E. Lemaire).

Let f be a complex-valued function defined in a neighborhood of $z^0 \in \mathbb{C}^n$. If there exists a

convergent power series P with center z^0 such that at every point of a neighborhood of z^0 the value off and the sum of P coincide, then f is called **analytic** at z^0 in the sense of Weierstrass, and P is the **Taylor expansion** of f at z^0 .

C. Differentiability

Let f be a complex-valued function defined in a neighborhood of $z^0 \in \mathbb{C}^n$. If in a neighborhood of z^0 we have

$$f(z) - f(z^0) = \alpha_1(z_1 - z_1^0) + \dots + \alpha_n(z_n - z_n^0) + \varepsilon, \\ \text{with } \alpha_1, \dots, \alpha_n \in \mathbb{C} \text{ and} \\ \lim_{z \rightarrow z^0} \varepsilon / (|z_1 - z_1^0| + \dots + |z_n - z_n^0|) = 0, \quad (1)$$

then we say that f is **(totally) differentiable** at z^0 . The function f is then continuous at z^0 , and the partial derivatives $\partial f / \partial z_j$ ($j = 1, \dots, n$) exist. Furthermore, the Cauchy-Riemann differential equations $\partial f / \partial \bar{z}_j = 0$ ($j = 1, \dots, n$) hold, where $\partial f / \partial z_j = (1/2)(\partial f / \partial x_j - i \partial f / \partial y_j)$ and $\partial f / \partial \bar{z}_j = (1/2)(\partial f / \partial x_j + i \partial f / \partial y_j)$ with $z_j = x_j + iy_j$. We say that f is **holomorphic** at z^0 in the sense of Riemann if f is differentiable at every point in a neighborhood of z^0 . Analyticity in the sense of Weierstrass is equivalent to holomorphy in the sense of Riemann. Furthermore, if the partial derivatives $\partial f / \partial z_j$ ($j = 1, \dots, n$) exist at every point in a neighborhood of z^0 , then f is, without assuming continuity, proved to be holomorphic. Thus the holomorphy off in each variable z_j implies the holomorphy off in $z = (z_1, \dots, z_n)$ (**Hartogs's theorem of holomorphy**, 1906).

A complex-valued function in a domain $G \subset \mathbb{C}^n$ is called **holomorphic** or **analytic** in G if it is holomorphic at every point of G . Let $H(G)$ be the ring of holomorphic functions in G . For $f = u + iv \in N(G)$, u and v satisfy in G the differential equations $\partial^2 u(z, \bar{z}) / \partial z_j \partial \bar{z}_k = 0$; that is,

$$\frac{\partial^2 u}{\partial x_j \partial x_k} + \frac{\partial^2 u}{\partial y_j \partial y_k} = 0, \quad (2) \\ \frac{\partial^2 u}{\partial x_j \partial y_k} - \frac{\partial^2 u}{\partial x_k \partial y_j} = 0, \quad j, k = 1, \dots, n.$$

A \dagger distribution $\text{TED}'(G)$ is called **pluriharmonic** in G if it satisfies (2) in G . Then T is harmonic and hence a real analytic function by \dagger Weyl's lemma.

Let G_j be a domain in the z_j -plane with piecewise smooth boundary C_j . If $f \in H(G)$ ($G = \prod_{j=1}^n G_j$) is continuous on \bar{G} , then

$$\frac{1}{(2\pi i)^n} \int_{C_1 \times \dots \times C_n} \frac{f(\zeta)}{(\zeta_1 - z_1) \dots (\zeta_n - z_n)} \\ d\zeta_1 \wedge \dots \wedge d\zeta_n = \begin{cases} f(z), & z \in G, \\ 0, & z \notin \bar{G} \end{cases} \quad (3)$$

(**Cauchy's integral representation**). Thus if $n \geq 2$, then f is determined by its values only on the proper subset $C = C_1 \times \dots \times C_n$ of ∂G , which is called the **skeleton** (or **determining set**) of G . For (pluriharmonic) functions of several complex variables, the boundary value problem, not necessarily solvable in its classical form, is not so effective as the Dirichlet problem in one complex variable.

As in the case of one variable, the \dagger Laurent expansion is valid for every holomorphic function in a domain of the form $G = \prod_{j=1}^n G_j$, where the G_j are circular annuli $\subset \mathbb{C}$. Suppose that we are given $f_1 \in H(G_1)$ and $f_2 \in H(G_2)$, where G_1, G_2 are domains in \mathbb{C}^n such that $G_1 \cap G_2$ is nonempty and connected. If $f_1 = f_2$ on $\{z \mid |z_j - z_j^0| < r_j, y = y^0, j = 1, \dots, n\}$, where $z^0 = x^0 + iy^0 \in G_1 \cap G_2$, then there exists a unique $f \in H(G_1 \cup G_2)$ such that $f|_{G_1} = f_1$ and $f|_{G_2} = f_2$ (**theorem of identity**). Thus \dagger analytic continuation proceeds as in the case of one variable. Similarly, some fundamental theorems in one variable, such as \dagger Liouville's theorem on \dagger entire functions and the \dagger maximum principle, hold also for several variables. However, there are some properties that reveal the differences between the cases of one and several variables. For instance, the set of zeros of a holomorphic function (\rightarrow 23 Analytic Spaces B) has no isolated point for $n \geq 2$. The investigation of these remarkable differences is one of the purposes of the theory of analytic functions of several complex variables.

D. Shilov Boundaries

While the maximum principle holds for a holomorphic function in a domain G , the set of points where the maximum is attained may be a proper subset S of ∂G . For instance, if G is the product of annuli as before, then the skeleton of G can be taken as S . In connection with the theory of \dagger normed rings, G. E. Shilov proved that there exists a unique smallest member S_0 (called the **Shilov boundary** of G) in the family of closed subsets S such that $\sup\{|f(z)| \mid z \in S\} = \sup\{|f(z)| \mid z \in G\}$ for every $f \in H(G)$ continuous on G . The structure of S_0 is investigated in detail together with the pseudoconvexity of G connected with it. Applying \dagger Perron's method for the Dirichlet problem to \dagger plurisubharmonic functions and the Shilov boundary, H. J. Bremermann solved one type of boundary problem.

E. Local Theory

Let f and g be two functions defined in neighborhoods of a set $S \subset \mathbb{C}^n$. If $f = g$ in a neighborhood of S , then f and g are called **equiva-**

Analytic Functions, Several Complex Variables

lent with respect to S . The germ off on S , denoted by f_S , is the equivalence class off: A **germ of a holomorphic function** on S is the germ on S of a holomorphic function defined in a neighborhood of S , and $H(S)$ denotes the ring of germs of holomorphic functions on S . Given a point 0 in \mathbb{C}^n , $H(0) = H(\{0\})$ is isomorphic to the ring H_n of convergent power series at 0 , i.e., the power series that are absolutely convergent in some neighborhoods of 0 . For every nonzero function $f \in H(0)$, there exists a system of coordinates (z_1, \dots, z_n) centered at 0 such that $f(0, \dots, z_n) \neq 0$ for every $z_n \neq 0$ in a neighborhood of $z_n = 0$. In a neighborhood of 0 , f is then equal to the product of an invertible element of H_n and a **distinguished pseudopolynomial**

$$P(z_n) = z_n^p + a_1(z_1, \dots, z_{n-1})z_n^{p-1} + \dots + a_p(z_1, \dots, z_{n-1}) \in H_{n-1}[z_n],$$

with $a_1(0, \dots, 0) = \dots = a_p(0, \dots, 0) = 0$, and $P(z_n)$ is uniquely determined by f and the coordinates z_1, \dots, z_n (**Weierstrass's preparation theorem**). It follows from this that H_n is an n -dimensional \dagger regular local ring. Considering $H(0)$ as the \dagger inductive limit of \dagger locally convex rings $H(U)$, where U ranges over a base for a neighborhood system of 0 , H. Cartan proved the preparation theorem in a more precise form in which the association $f \rightarrow a_j$ is continuous with respect to the supremum norm. Based on a deep consideration of this situation, K. Oka proved a theorem of fundamental importance: The \dagger sheaf $\mathcal{O}_{\mathbb{C}^n}$ defined by $\mathcal{O}_{\mathbb{C}^n, z} = H(z) (z \in \mathbb{C}^n)$ is \dagger coherent.

F. Domains of Holomorphy

Given a domain $G \subset \mathbb{C}^n$ for $n \geq 2$, it may be that there exists a domain G' strictly greater than G such that all the functions that are holomorphic in G extend to holomorphic functions in G' . For instance, let $S = S' \times \sigma$, where S' and σ are open polydisks in (z_1, \dots, z_{n-1}) -space and z_n -space, respectively, and let $T \subset \mathbb{C}^n$ be an open set. If there exists an open set $U (\neq \emptyset) \subset S'$ such that $(U \times \sigma) \cup (S' \times \partial\sigma) \subset T$ and if $S \cap T$ is connected, then all the functions that are holomorphic in T extend uniquely to holomorphic functions in $S \cup T$ (**Hartogs's continuation theorem**). In particular, if A is an \dagger analytic set in a domain $G \subset \mathbb{C}^n$ with $\dim A \leq n - 2$, then all the functions that are holomorphic in $G - A$ extend uniquely to holomorphic functions in G . Furthermore, if A is an analytic set in G with $A \neq G$, then every $f \in H(G - A)$ that is locally bounded at the points of A extends uniquely to a holomorphic function in G (**Riemann's continuation theorem**

for $n \geq 2$). The domain \tilde{G}_f of holomorphy for f is defined to be the maximal domain to which f may be continued analytically. A domain G is called a **domain of holomorphy** if $G = \tilde{G}_f$ for some $f \in H(G)$. However, \tilde{G}_f is, in general, not a subdomain of \mathbb{C}^n . \tilde{G}_f is, generally, a manifold spread over \mathbb{C}^n ; i.e., \tilde{G}_f is a connected n -dimensional \dagger complex analytic manifold with a holomorphic mapping $\varphi: \tilde{G}_f \rightarrow \mathbb{C}^n$ of maximum Jacobian rank (φ is then an open mapping). The same is true for the common existence domain of functions in a subfamily of $H(G)$. The common existence domain \tilde{G} of all the functions in $H(G)$ is called the **envelope of holomorphy**. A **holomorphically complete domain** is a domain G such that $G = \tilde{G}$. These notions carry over to the case where G is a manifold spread over \mathbb{C}^n . The (general) Levi problem of determining the conditions for a given domain to be holomorphically complete is fundamental to the theory of analytic functions of several complex variables (\rightarrow Section I). In connection with this problem, various notions of pseudoconvexity of holomorphically complete domains are defined.

G. Pseudoconvexity

An upper semicontinuous real-valued function $u (-\infty \leq u < +\infty)$ in a domain $G \subset \mathbb{C}^n$ is said to be **plurisubharmonic** if for every $z^0 \in G$ and every $a \in \mathbb{C}^n$ the function $u(z^0 + ta)$ of t is \dagger subharmonic (including the constant $-\infty$) in all the connected components of $\{t \mid z^0 + ta \in G\}$. A domain G is said to be **pseudoconvex** (or **d-pseudoconvex**) if $u = -\log d$, is plurisubharmonic in G , where $d(z)$ is the distance from $z \in G$ to ∂G with respect to any norm in \mathbb{C}^n . Every connected component of the interior of the intersection of a family of pseudoconvex domains is pseudoconvex, and the union of an increasing sequence of pseudoconvex domains is pseudoconvex. Suppose that we are given a domain G and a function u of class C^2 in a neighborhood of \bar{G} such that $G = \{z \mid u(z) < 0\}$ and, for some $\varepsilon > 0$, $\sum_{j,k} (\partial^2 u / \partial z_j \partial \bar{z}_k) a_j \bar{a}_k \geq \varepsilon |a|^2$ for every $a \in \mathbb{C}^n$. Then the domain G is said to be **strongly pseudoconvex**. Strong pseudoconvexity implies pseudoconvexity. Every pseudoconvex domain is exhausted by an increasing sequence of strongly pseudoconvex domains. An open set $P \subset \mathbb{C}^n$ is called an **analytic polyhedron** if $P = \{z \mid |\chi_\alpha(z)| < 1, \alpha = 1, \dots, N\}$, $\chi_\alpha \in H(\bar{P})$ ($\alpha = 1, \dots, N$). Then every connected component of P is pseudoconvex. A **Weil domain** is a connected and bounded analytic polyhedron P defined by $\chi_\alpha (\alpha = 1, \dots, N)$ with $N \geq n$, such that for every k ($1 \leq k \leq n$) the intersection of the hypersurfaces $|\chi_{\alpha_i}(z)| = 1$ ($1 \leq i \leq k$) is of dimension $\leq 2n - k$.

H. Holomorphic Convexity

A domain $G \subset \mathbb{C}^n$ is called **holomorphically convex** if for every compact set $K \subset G$, $\hat{K} = \bigcap_{f \in H(G)} \{z \mid |f(z)| \leq \sup_{\omega \in K} |f(\omega)|\}$ (the **holomorphic hull** of K) is a compact set contained in G . (For a domain G contained in an \dagger analytic set we can similarly define holomorphic convexity of G .) Every connected component of the intersection of a family of holomorphically convex domains is holomorphically convex, and a holomorphically convex domain is exhausted by an increasing sequence of Weil domains. Holomorphic completeness implies holomorphic convexity. The converse is true for domains in \mathbb{C}^n . If G is holomorphically convex, then for every point ζ of ∂G there exists an $f \in H(G)$ such that f is not locally bounded at ζ (H. Cartan and P. Thullen, 1932). Hence a holomorphically convex domain is a domain of holomorphy. Thus a domain in \mathbb{C} is holomorphically convex if and only if it is a domain of holomorphy. (The same is true for unramified covering domains over \mathbb{C}^n (Oka, 1953).) The union of an increasing sequence of domains of holomorphy is a domain of holomorphy (H. Behnke and K. Stein [2], 1939). Suppose that we are given a domain G and domains S_α , T_α ($\alpha = 1, 2, \dots$) such that $S_\alpha \cup T_\alpha \subset G$ and $\sup_{T_\alpha} |f| = \sup_{S_\alpha \cup T_\alpha} |f|$ for every $f \in H(G)$. Suppose also that $S_0 = \lim S_\alpha$ is bounded. We say that the **continuity principle** holds in G if $\bar{T}_0 \subset G$ ($T_0 = \lim T_\alpha$) implies $S_0 \subset G$. The continuity principle holds in a domain of holomorphy (**Hartogs's theorem of continuity**). This implies that if G is a bounded domain $\subset \mathbb{C}^n$ ($n \geq 2$) with connected boundary ∂G , then every function holomorphic in a neighborhood of ∂G extends to a holomorphic function in G (**Hartogs-Osgood theorem**). In particular, for $n \geq 2$, the set of singular points of a holomorphic function has no isolated point. A domain is pseudoconvex if the continuity principle holds there. Hence a domain of holomorphy is pseudoconvex.

I. The Levi Problem

Let G be a domain in \mathbb{C}^n and $z^0 \in \partial G$. If there exists an open neighborhood U of z^0 such that every connected component of $G \cap U$ is a domain of holomorphy, then G is called **Cartan pseudoconvex** at z^0 . On the other hand, if every 1-dimensional analytic set that has z^0 as an ordinary point contains points not belonging to $G \cup \{z^0\}$ in the neighborhoods of z^0 , then G is called **Levi pseudoconvex** at z^0 . Furthermore, G is called **locally Cartan (Levi) pseudoconvex** if G is Cartan (Levi) pseudoconvex at every point of ∂G . Every domain of

holomorphy is locally Cartan pseudoconvex. If G is pseudoconvex and there exists a neighborhood U of z^0 such that $G \cap U = \{z \mid \varphi(z) < 0\}$, where $\varphi \in C^1(U)$, then G is Levi pseudoconvex at z^0 .

The (proper) **Levi problem** of whether every pseudoconvex domain is a domain of holomorphy was proposed by E. E. Levi (1911). After unsuccessful efforts by various mathematicians to solve the problem, it was affirmatively solved by Oka (1942 for $n = 2$ and 1953 for manifolds spread over \mathbb{C}^n for $n \geq 2$), H. Bremermann, and F. Norguet. The problem was solved also by H. Grauert [14] (1958) in a more general form (- Section L) using results on linear topological spaces and by L. Hörmander [18] (1965) using methods of the theory of partial differential equations. A fundamental step in Oka's solution is his **gluing theorem**: Let G be a bounded domain $\subset \mathbb{C}^n$. If every connected component of $G_1 = \{z \mid x_1 > a\} \cap G$ and $G_2 = \{z \mid x_1 < b\} \cap G$ ($a < b$) is a domain of holomorphy, then G is a domain of holomorphy. Indeed, by virtue of the Behnke-Stein theorem and the fact that every pseudoconvex domain is the union of an increasing sequence of bounded locally Cartan pseudoconvex domains, it suffices to solve the Levi problem in the case of a bounded locally Cartan pseudoconvex domain. The Levi problem in this case is solved by the gluing theorem. Various integral representations of holomorphic functions are known besides the Cauchy representation. The Bergmann-Weil integral representation in a Weil domain was used as an important means of solving the Levi problem.

J. Holomorphic Mappings

Holomorphic functions with values in a \dagger quasi-complete \dagger locally convex complex vector space E have also been investigated. The classical theory described above has been generalized, to some extent, to this case. In this way, many applications of the theory have been discovered. An E -valued function in a domain $G \subset \mathbb{C}^n$ is holomorphic if and only if the mapping $u \mapsto \langle u, f \rangle$ is holomorphic for every continuous linear form u on E . By this theorem, most problems concerning E -valued holomorphic functions can be reduced to those of ordinary holomorphic functions. Note that the vector space $H(G)$ of ordinary holomorphic functions in G is a \dagger Fréchet space. The spaces \mathbb{C}^p and complex \dagger Banach spaces belong to the above category of E . A \mathbb{C}^p -valued holomorphic function in a domain $G \subset \mathbb{C}^n$ is called a **holomorphic (or analytic) mapping** of G into \mathbb{C}^p . An isomorphism in the category of do-

mains $G \subset \mathbb{C}^n$ and holomorphic mappings is called an **analytic isomorphism** (or **biholomorphic mapping**). An automorphism in the category is called an **analytic** (or **holomorphic**) **automorphism**. With every domain $G \subset \mathbb{C}^n$ is associated the sheaf \mathcal{O}_G of germs of holomorphic functions over G . Thus we have the notion of a **ringed space** (G, \mathcal{O}_G) . A complex analytic manifold can be defined as a (Hausdorff) ringed space that is locally isomorphic to some (G, \mathcal{O}_G) .

A **meromorphic function** in G is a function that is locally the quotient of two holomorphic functions with denominator $\neq 0$. It may be defined more rigorously as a meromorphic mapping of G into $\mathbb{P}_n(\mathbb{C})$ (- 23 Analytic Spaces D).

K. The Cousin Problems

The Cousin problems are those of constructing meromorphic functions with given **zeros** or **poles**. In terms of sheaves the problems are stated as follows: Let \mathcal{K}_G be the sheaf of germs of meromorphic functions over a domain $G \subset \mathbb{C}^n$. The **first Cousin problem** asks whether the mapping $\Gamma(G, \mathcal{K}_G) \rightarrow \Gamma(G, \mathcal{P}_G)$ induced by the exact sequence $0 \rightarrow \mathcal{O}_G \rightarrow \mathcal{K}_G \rightarrow \mathcal{P}_G \rightarrow 0$ ($\mathcal{P}_G = \mathcal{K}_G / \mathcal{O}_G$) is surjective, where $\Gamma(G, \mathcal{F})$ is the module of **sections** of \mathcal{F} over G (- 383 Sheaves C). Let \mathcal{K}_G^* be the sheaf of multiplicative groups of germs of meromorphic functions not identically 0 and \mathcal{O}_G^* be the subsheaf of \mathcal{K}_G^* formed by germs of nonzero holomorphic functions. The **second Cousin problem** asks whether the mapping $\Gamma(G, \mathcal{K}_G^*) \rightarrow \Gamma(G, \mathcal{D}_G)$ ($\mathcal{D}_G = \mathcal{K}_G^* / \mathcal{O}_G^*$) is surjective. P. Cousin (1895) solved the first problem for $G = \mathbb{C}^n$ or $\prod_{j=1}^n G_j$ and the second problem for $G = \mathbb{C}^n$. Oka (1935) proved that the first problem is solvable in every domain of holomorphy. In solving the second problem in a domain of holomorphy, Oka established the notion of **fiber bundles** and proved that the problem for any domain is reduced to holomorphic triviality of a holomorphic principal fiber bundle over the domain and that holomorphic triviality is equivalent to topological triviality when the domain is of holomorphy (**Oka's principle**). Using the solutions of the Cousin problems, Oka proved his gluing theorem, described in Section 1.

L. Stein Manifolds

Abstracting certain important properties of a domain of holomorphy, Stein [33] introduced the following category of complex analytic manifolds (X, \mathcal{O}_X) : (1) X is paracompact (i.e.,

each connected component of X has a countable open base). (2) Functions in $\Gamma(X, \mathcal{O}_X)$ separate the points of X . (3) For every point $x \in X$ there exists a system of local coordinates around x that is formed by functions in $\Gamma(X, \mathcal{O}_X)$. (4) X is holomorphically convex. (X, \mathcal{O}_X) is then called a **Stein manifold**. It was later discovered by Grauert [12] that conditions (2) and (4) imply (1) and (3).

Applying the theory of cohomology with coefficients in sheaves, H. Cartan and J.-P. Serre obtained for an analytic coherent sheaf \mathcal{F} on a Stein manifold X , the following **fundamental theorems of Stein manifolds**. **Theorem A:** $H^0(X, \mathcal{F})$ generates the stalk \mathcal{F}_x (as an \mathcal{O}_x -module) at every point x of X . **Theorem B:** $H^q(X, \mathcal{F}) = 0$ for all $q > 0$ [7, 30]. Conversely, for a complex analytic manifold X , if for every analytic coherent sheaf \mathcal{F} of ideals defined by a 0-dimensional analytic set in X (i.e., a discrete subset of X), $H^1(X, \mathcal{F}) = 0$, then X is a Stein manifold. Furthermore, if $\Gamma(X, \mathcal{O}_X) = \Gamma(Y, \mathcal{O}_Y)$ for a Stein manifold Y (as in the case where $X \subset \mathbb{C}^n$), then the fundamental theorem A for every coherent sheaf of ideals implies that X is a Stein manifold (I. Wakabayashi).

Due to the fundamental theorems, most results on domains of holomorphy hold unchanged for Stein manifolds. For instance, the first Cousin problem is always solvable. The second Cousin problem is solvable if and only if $H^2(X, \mathbb{Z}) = 0$. An n -dimensional Stein manifold can be realized as a (ramified) covering domain of holomorphy over \mathbb{C}^n . Furthermore, some theorems on differentiable manifolds have analogs on Stein manifolds. For instance, the cohomology groups of the complex of holomorphic differential forms over a Stein manifold X are isomorphic to the cohomology groups $H^*(X, \mathbb{C})$ (analog of **de Rham's theorem**). Every n -dimensional Stein manifold X is realized as a closed complex analytic submanifold in \mathbb{C}^{2n+1} ; that is, there exists an injective **proper** holomorphic mapping $f: X \rightarrow \mathbb{C}^{2n+1}$ with $df \neq 0$. Consider all the holomorphic **principal fiber bundles** over a Stein manifold X whose fibers are isomorphic to a complex Lie group G . The analytic isomorphism classes of the bundles and the elements in $H^1(X, G^a)$ (where G^a is the sheaf of germs of holomorphic mappings of X into G) are in one-to-one correspondence. The same is true for the topological isomorphism classes of the bundles and the elements in $H^1(X, G^c)$ (where G^c is the sheaf of germs of continuous mappings of X into G). The mapping $H^1(X, G^a) \rightarrow H^1(X, G^c)$ induced by the canonical injection $G^a \rightarrow G^c$ is bijective (Grauert [13]). Every relatively compact domain in a complex analytic manifold is holomorphically convex if it is strongly pseudoconvex (Grauert [14]). Hence

such a domain is a Stein manifold. It follows from this that every real analytic manifold with countable base for open sets is realized as a closed real analytic submanifold of some \mathbf{R}^n .

The notion of Stein manifolds is generalized to that of weakly 1-complete manifolds: a complex manifold X is called a **weakly 1-complete manifold** if there exists a plurisubharmonic function u of class C^∞ on X , such that for any $c \in \mathbf{R}$, $X_c = \{x \in X \mid u(x) < c\}$ is relatively compact. The family of such manifolds includes compact complex manifolds too. For a weakly 1-complete manifold also, vanishing theorems of cohomology have been established by H. Hironaka, S. Nakano [26], and H. Kazama [2] (- 232 Kähler Manifolds D).

The theory of entire functions of two variables has been developed from a new viewpoint established by T. Nishino. An entire function f of two variables defines a family of Riemann surfaces $\{f=c \mid c \in \mathbf{C}\}$ on \mathbf{C}^2 , and investigations of the structure of such a family play an important role. For instance the following is proved in this way: If every irreducible component of $f=c$ ($c \in \mathbf{C}$) is biholomorphic to \mathbf{C}^1 , then there exists an entire function g such that $(f, y): \mathbf{C}^2 \rightarrow \mathbf{C}^2$ is a biholomorphic mapping (Nishino [28]). In this theorem, the analyticity of g is obtained from the fact that \mathbf{C}^2 is a Stein manifold.

M. Continuation of Analytic Sets

The application of the theory of cohomology with coefficients in sheaves is not restricted to problems concerning Stein manifolds. Given $G_0 = \{z \mid |z_j| < 1, 1 \leq j \leq n\}$ ($n \geq 3$), $G_1 = \{z \mid |z_1| < \varepsilon, |z_j| < 1, 2 \leq j \leq n\}$, and $G^{(m)} = G_1 \cup (G_0 - \{z \mid z_2 = \dots = z_m = 0\})$ ($3 \leq m \leq n$), we have $H^p(G^{(m)}, \mathcal{O}_{G^{(m)}}) = 0$ ($1 \leq p \leq m-2$) (Scheja's theorem [32]). Let \mathcal{F} be a coherent analytic sheaf over a domain $G \subset \mathbf{C}^n$. If, for every point z of an analytic set $A \subsetneq G$, $\mathcal{F}_z = \{0\}$ or $p \leq n - \dim A - 2$ hd, \mathcal{F} (where hd, \mathcal{F} is the homological dimension of the $\mathcal{O}_{G,z}$ -module \mathcal{F}_z), then it follows from Scheja's theorem that the mapping $H^p(G, \mathcal{F}) \rightarrow H^p(G - A, \mathcal{F})$ induced by the canonical injection $G - A \rightarrow G$ is bijective. This generalizes Hartogs's continuation theorem for holomorphic functions, which corresponds to the case $p=0$.

Besides the continuation of holomorphic functions, we can consider the continuation of analytic sets. Let A be an analytic set in a domain $G \subset \mathbf{C}^n$ and S an analytic set in $G - A$. A point $z \in G$ is said to be **regular (essentially singular)** with respect to S if the closure \bar{S} of S in G is (is not) analytic at z . If $\dim A < \dim S$ for every point $z \in S$, then \bar{S} is analytic in G . If S is purely d -dimensional and S is analytic at

a point of a d -dimensional irreducible component A' of A , then S is analytic at every point of A' that is not located in any other irreducible component of A . Furthermore, if $\dim A \leq \dim S$ and S is purely d -dimensional, then the following hold: (1) The set E of essential singularities of S is, if not empty, a purely d -dimensional analytic set in G , formed by irreducible components of A . (2) If every irreducible component of A contains points of E not located in any other irreducible component of A , then $A \subset E$, and A is, if not empty, purely d -dimensional. (3) If every d -dimensional irreducible component of A contains points that are regular with respect to S , then S is a purely d -dimensional analytic set in G (Thullen, Remmert, and Stein). By these results it is possible to give a proof for †Chow's theorem that every analytic set in $\mathbf{P}_n(\mathbf{C})$ is algebraic.

The continuation of holomorphic functions is related to the continuation of their graphs. W. Rothstein investigated the continuation of analytic sets to obtain the following analog to Hartogs's theorem of continuity: If $G = G_1 \cup G_2$, $G_1 = \{z \mid |z_1| < 1/2, \sum_{j=2}^n |z_j|^2 < 1\}$, $G_2 = \{z \mid |z_1| < 1, 1/2 < \sum_{j=2}^n |z_j|^2 < 1\}$, and $\tilde{G} = \{z \mid |z_1| < 1, \sum_{j=2}^n |z_j|^2 < 1\}$ (the envelope of holomorphy of G) with $n \geq 3$, then every purely $(n-1)$ -dimensional analytic set A in G extends to an analytic set in \tilde{G} ; that is, there exists a purely $(n-1)$ -dimensional analytic set \tilde{A} in \tilde{G} such that $A = \tilde{A} \cap G$. K. Kasahara and H. Fujimoto generalized this theorem to the case of analytic spaces.

N. Nevanlinna Theory for Several Complex Variables

The Nevanlinna theory investigates holomorphic mappings between complex manifolds. In function theory of one variable, for a holomorphic mapping $f: \mathbf{C} \rightarrow \mathbf{P}^1(\mathbf{C})$ from \mathbf{C} into the 1-dimensional †complex projective space, the famous †Picard theorem states that if f omits three values, then f must be constant. R. Nevanlinna developed the quantitative theory of value distributions off: L. Ahlfors [1], introduced the geometric approach, and enunciated the principle that the negative curvature of the image manifold restricts a holomorphic mapping. In higher-dimensional situations, this principle has been realized in many cases. The Nevanlinna theory for a holomorphic mapping $f: \mathbf{C} \rightarrow \mathbf{P}^n(\mathbf{C})$ with respect to †hyperplanes was established by Ahlfors and H. Weyl and J. Weyl [34]. Holomorphic mappings into projective spaces have been studied in detail by W. Stoll, H. Fujimoto, and M. Green. Following the work of S. S.

Chern [9] of around 1970, the equidimensional case was investigated (P. A. Griffiths, K. Kodaira, S. Kobayashi and T. Ochiai). J. A. Carlson and Griffiths [6] succeeded in obtaining a defect relation for equidimensional holomorphic mappings and hypersurfaces of the image manifold. This yields the following Picard-type theorem (after the formulation by F. Sakai, *Inventiones Math.*, 16 (1974)): Let X be a projective algebraic manifold of dimension n and $f: \mathbb{C}^n \rightarrow X$ a holomorphic mapping. Let D be a hypersurface of X . We take a desingularization $\pi: X^* \rightarrow X$ so that the inverse image $D^* = \pi^{-1}(D)$ has at most normal crossings as singularities. We denote by K^* a canonical divisor of X^* . We assume that

$$\limsup_{m \rightarrow \infty} \frac{\dim H^0(X^*, \mathcal{O}(m(K^* + D^*)))}{m^n} > 0.$$

If f omits D , then the Jacobian of f vanishes everywhere. As a corollary, it follows that the universal covering manifold of $X - D$ cannot be \mathbb{C}^n . When $X = \mathbb{P}^n(\mathbb{C})$ and D is a nonsingular hypersurface of degree d , the above assumption is satisfied if $d \geq n + 2$. In the nonequidimensional case, Ochiai [29] (with supplementary works by M. Green, Y. Kawamata, and P. Wong) verified the following assertion, which was first stated by A. Bloch with rough arguments: Let X be a projective algebraic manifold of dimension n , and suppose $\dim H^1(X, \mathcal{O}_X) > n$; then the image of a holomorphic mapping $f: \mathbb{C} \rightarrow X$ is contained in a proper subvariety of X (\rightarrow 124 Distributions of Values of Functions of a Complex Variable; 272 Meromorphic Functions).

0. Hyperbolic Manifolds

Every complex analytic space X has two invariant pseudodistances: the Carathéodory pseudodistance c_X and the Kobayashi pseudodistance d_X , both of which generalize the Poincaré distance ρ of the unit disk $D = \{ |z| < 1 \}$. These pseudodistances can be defined by the property that d_X is the largest pseudodistance among all pseudodistances δ_X on X for which all holomorphic mappings $(X, \delta_X) \rightarrow (D, \rho)$ are distance-decreasing, while c_X is the smallest among all pseudodistances δ_X for which all holomorphic mappings $(D, \rho) \rightarrow (X, \delta_X)$ are distance-decreasing. Then $c_X \leq d_X$. If Y is another complex analytic space, every holomorphic mapping $f: X \rightarrow Y$ is distance-decreasing with respect to either its Carathéodory or Kobayashi pseudodistance. This may be considered to be a generalization of the Schwarz-Pick lemma. (The Schwarz lemma is often

referred to by this name, since G. Pick also investigated the distance-decreasing property.) The Kobayashi pseudodistance d_X can also be obtained by integrating an infinitesimal differential pseudometric called the Kobayashi pseudometric in the same manner as the Riemannian distance is obtained from the Riemannian metric (H. L. Royden).

If d_X is a (complete) distance on X , then X is said to be (complete) hyperbolic. A Riemann surface is hyperbolic in this sense if and only if its universal covering is biholomorphic to the unit disk. If X is open in Y and if, for every pair of sequences of points $\{p_n\}$ and $\{q_n\}$ in X converging to distinct points of X , $\lim d_X(p_n, q_n)$ is positive, then X is said to be hyperbolically embedded in Y . The Riemann sphere $\mathbb{C}U \setminus \{\infty\}$ minus three points is not only complete hyperbolic but also hyperbolically embedded in the Riemann sphere. Every holomorphic mapping of \mathbb{C} into a hyperbolic complex analytic space is constant, while every holomorphic mapping of the punctured disk $D^* = \{0 < |z| < 1\}$ into a hyperbolically embedded space $X \subset Y$ extends to a holomorphic mapping of D into Y . Thus the classical little Picard theorem reduces to the statement that $\mathbb{C} \setminus \{0, 1\}$ is hyperbolic, while the great Picard theorem reduces to showing that $\mathbb{C} \setminus \{0, 1\}$ is hyperbolically embedded in the Riemann sphere. These classical theorems can be generalized in two ways. If M is a symmetric bounded domain and Γ is a discrete arithmetic group acting freely on M , then M/Γ is not only complete hyperbolic but also hyperbolically embedded in its Satake compactification (Kobayashi and Ochiai). If X is the complement of $2n + 1$ hyperplanes in general position in the complex projective space $\mathbb{P}^n(\mathbb{C})$, then X is complete hyperbolic and hyperbolically embedded in $\mathbb{P}^n(\mathbb{C})$ (a restatement of a result going back to E. Borel, A. Bloch, and H. Cartan).

Although there are some noncompact nonhyperbolic complex manifolds X for which every holomorphic mapping $f: \mathbb{C} \rightarrow X$ is constant, a compact complex manifold X is hyperbolic if and only if every holomorphic mapping $f: \mathbb{C} \rightarrow X$ is constant (R. Brody).

For the Teichmüller space $X = \mathbf{T}_g$ of compact Riemann surfaces of genus g , the Kobayashi distance d_X agrees with the Teichmüller distance which had been introduced before the complex structure of \mathbf{T}_g was defined (H. L. Royden).

In the study of pseudoconvex domains, the Bergman metric, the Carathéodory distance, and the Kobayashi distance serve as useful tools. Their behavior at the boundary has been studied extensively.

Given an n -dimensional complex manifold X , an invariant measure Ψ_X can be defined as the largest measure such that every holomorphic mapping f from the polydisk D^n with invariant measure into X is measure-decreasing. For an algebraic manifold of general type, this measure is everywhere positive.

P. Bounded Domains in \mathbb{C}^n

For any bounded domain in \mathbb{C}^n there is naturally assigned a †Kähler metric called the Bergman metric. Using the invariance of this metric, E. Cartan proved that all the Hermitian symmetric spaces of noncompact type are realized as bounded domains. In view of the fact that some important period matrix domains (e.g., the 19-dimensional bounded symmetric domain of type IV, the Siegel upper space, etc.) are of this type, it is obviously of great significance in algebraic geometry to study discontinuous subgroups of the automorphism groups of such domains (Pyatetskii-Shapiro).

On the other hand, it is almost impossible for general bounded domains to determine the explicit form of their Bergman metrics. But Cartan also investigated strongly pseudoconvex real hypersurfaces in \mathbb{C}^2 , and he solved completely the (local) equivalence problem for them, introducing a definite type of †Cartan connections over them in a functorial way. Thereafter N. Tanaka (and for hypersurfaces also Chern and J. Moser) generalized this result to all pseudoconvex real submanifolds in higher-dimensional complex manifolds. One can further apply this result to the equivalence problem of bounded domains with smooth strongly pseudoconvex boundaries, for C. Fefferman [10] proved by analyzing the boundary behavior of the Bergman metric that every biholomorphic mapping between two such domains is extended smoothly up to their boundaries (I. Naruki [27] gave an alternative proof).

Q. History

In connection with †Abelian functions, analytic functions of several complex variables have been studied sporadically since the time of Riemann and Weierstrass (H. Poincaré, Cousin). A series of investigations by Hartogs ([17] (1906), etc.) that revealed the distinctive properties of several complex variables initiated a new epoch in complex analysis. Levi (1910–1911) generalized Hartogs's results to the case of meromorphic functions, introduced the notion of pseudoconvexity, and proposed

the so-called Levi problem. After a lapse of time, many contributions to this new area of complex analysis have been made since 1920. The study by K. Reinhardt (1921) of analytic automorphisms was further developed by C. Carathéodory and Behnke. The †kernel function introduced by S. Bochner and S. Bergmann (1922) produced many remarkable results. In contrast with †Picard's theorem in one variable, P. Fatou found a holomorphic mapping $f: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ with nonvanishing Jacobian such that the image $f(\mathbb{C}^2)$ has an exterior point.

The theory of analytic functions of several complex variables has flourished since 1926. Behnke and Thullen in Münster, together with G. Julia and H. Cartan in Paris, were the most active investigators. The results on †normal families of analytic functions of several complex variables (Julia [20], 1926), the uniqueness theorem of holomorphic mappings (H. Cartan, 1930), and a characterization of a domain of holomorphy by holomorphic convexity (Cartan and Thullen, 1932) are their most remarkable achievements. Behnke and Thullen [3] systematized the results obtained since the discovery of the theory by providing a complete bibliography of articles up to 1934.

The three major unsolved problems at that time—those of Cousin, Levi, and the approximation of holomorphic functions—were intensively studied by Oka from 1936, who has given complete solutions [30]. The investigation of ideals of holomorphic functions by H. Cartan (1944), together with that of ideals with undetermined domains by Oka, has developed into the theory of coherent analytic sheaves. The notion of analytic spaces, first introduced by Behnke and Stein (1951), extended the field of investigation in the theory of analytic functions of several complex variables. The theory of cohomology with coefficients in sheaves has been effectively applied by H. Cartan and Serre (1951–1952). The introduction of the notion of Stein manifolds (1951) came at the same time. Grauert's deep investigations since 1955, together with those of Stein and Remmert, have contributed greatly to the development of the theory of analytic spaces. In the 1960s, active investigations took place also in the United States [16]. The theory of automorphic functions of several complex variables has been developed by C. L. Siegel, I. Satake, and others in connection with the theory of numbers. Entire functions of two variables have been investigated from a new point of view by Nishino and others since 1968 (→ last paragraph of Section N). The notion of a hyperbolic manifold introduced by Kobayashi (1967) enables us to obtain many results on

complex manifolds by means of the methods of differential geometry. The theory of automorphisms of bounded pseudoconvex domains was developed extensively in the 1970s by Fefferman, Naruki, and others. The theory of an envelope of holomorphy has also been successfully applied to the theory of elementary particles in physics.

References

- [1] L. Ahlfors, An extension of Schwarz's lemma, *Trans. Amer. Math. Soc.*, 43 (1938), 359–364.
- [2] H. Behnke and K. Stein, Konvergenzfolgen von Regularitätsbereichen und der Meromorphiekonvexität, *Math. Ann.*, 116 (1939), 204–216.
- [3] H. Behnke and P. Thullen, *Theorie der Funktionen mehrerer komplexer Veränderlichen*, Springer, 1934; second edition, 1970.
- [4] L. Bers, Introduction to several complex variables, *Courant Institute*, 1964.
- [5] S. Bochner and W. T. Martin, *Several complex variables*, Princeton Univ. Press, 1948.
- [6] J. A. Carlson and P. A. Griffiths, A defect relation for equidimensional holomorphic mappings between algebraic varieties, *Ann. Math.*, (2) 95 (1972), 557–584.
- [7] Sém. H. Cartan, *Ecole Norm. Sup.*, 1951–1952.
- [8] H. Cartan, *Oeuvres*, Springer, 1979.
- [9] S. S. Chern, On holomorphic mappings of Hermitian manifolds of the same dimension, *Amer. Math. Soc. Proc. Symposia in Pure Math.*, 11 (1968), 157–170.
- [10] C. Fefferman, The Bergman kernel and biholomorphic mappings of pseudoconvex domains, *Inventiones Math.*, 26 (1974), 1–65.
- [11] B. A. Fuks, Special chapters in the theory of analytic functions of several complex variables, *Amer. Math. Soc. Transl. of Math. Monographs*, 1965.
- [12] H. Grauert, Charakterisierung der holomorph-vollständiger komplexen Räume, *Math. Ann.*, 129 (1955), 233–259.
- [13] H. Grauert, Holomorphe Funktionen mit Werten in komplexen Lieschen Gruppen, *Math. Ann.*, 133 (1957), 450–472.
- [14] H. Grauert, On Levi's problem and the imbedding of real-analytic manifolds, *Ann. Math.*, (2) 68 (1958), 460–472.
- [15] H. Grauert and R. Remmert, *Theory of Stein spaces*, Springer, 1979.
- [16] R. C. Gunning and H. Rossi, *Analytic functions of several complex variables*, Prentice-Hall, 1965.
- [17] F. Hartogs, Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlichen, insbesondere über die Darstellung derselben durch Reihen, welche nach Potenzen einer Veränderlichen fortschreiten, *Math. Ann.*, 62 (1906), 1–88.
- [18] L. Hörmander, L^2 estimates and existence theorems for the $\bar{\partial}$ operator, *Acta Math.*, 113 (1965), 89–152.
- [19] L. Hörmander, An introduction to complex analysis in several complex variables, *Van Nostrand*, 1966.
- [20] G. Julia, Sur les familles de fonctions analytiques de plusieurs variables, *Acta Math.*, 47 (1926), 53–115.
- [21] H. Kazama, Approximation theorem and application to Nakano's vanishing theorem for weakly 1-complete manifolds, *Mem. Fac. Sci. Kyushu Univ.*, 27 (1973), 221–240.
- [22] S. Kobayashi, *Hyperbolic manifolds and holomorphic mappings*, Dekker, 1970.
- [23] S. Kobayashi, Intrinsic distances, measures and geometric function theory, *Bull. Amer. Math. Soc.*, 82 (1976), 357–416.
- [24] P. Lelong, *Fonctions plurisousharmoniques et formes différentielles positives*, Cordon & Breach, 1968.
- [25] L. Nachbin, *Holomorphic functions, domains of holomorphy and local properties*, North-Holland, 1970.
- [26] S. Nakano, Vanishing theorems for weakly 1-complete manifolds, II, *Publ. Res. Inst. Math. Sci.*, 10 (1974), 101–110.
- [27] I. Naruki, On extendibility of isomorphisms of Cartan connections and biholomorphic mappings of bounded domains, *Tôhoku Math. J.*, 28 (1976), 117–122.
- [28] T. Nishino, Nouvelles recherches sur les fonctions entières de plusieurs variables complexes (II), *Fonctions entières qui se réduisent à celles d'une variables*, *J. Math. Kyoto Univ.*, 9 (1969), 221–274.
- [29] T. Ochiai, On holomorphic curves in algebraic varieties with ample irregularity, *Inventiones Math.*, 43 (1977), 83–96.
- [30] K. Oka, *Sur les fonctions analytiques de plusieurs variables (collected papers)*, Iwanami, 1961.
- [31] W. F. Osgood, *Lehrbuch der Funktionentheorie II*, Teubner, 1924; revised edition, 1929 (Chelsea, 1965).
- [32] G. Scheja, Riemannsche Hebbarkeitssätze für Cohomologieklassen, *Math. Ann.*, 144 (1961), 345–360.
- [33] K. Stein, Analytische Funktionen mehrerer komplexer Veränderlichen zu vorgegebenen Periodizitätsmoduln und das zweite Cousinsche Problem, *Math. Ann.*, 123 (1951), 201–222.
- [34] H. Weyl and J. Weyl, Meromorphic functions and analytic curves, *Ann. Math.*

Studies 12, Princeton Univ. Press, 1943.
Also \rightarrow references to 23 Analytic Spaces.

22 (1.12) Analytic Sets

A. General Remarks

The notion of analytic sets was first defined by N. N. Luzin and M. Ya. Suslin in 1916, and it was extended to that of projective sets by operations such as complementation and projection (Luzin, 1924). Most mathematicians, including Luzin and W. Sierpinski, who worked in this field, were in agreement with †French empiricism (or tsemi-intuitionism), which defended the standpoint of R. Baire, E. Borel, H. Lebesgue, and others. An object is said to be **effectively given** if it can be uniquely, individually, and unambiguously determined in finite terms so that anyone can reach the same object by following the defining procedure. Semi-intuitionists claim that only effectively given objects have mathematical existence, and they do not recognize as a mathematical object something that needs the axiom of choice for its definition. From this point of view, †Borel sets were “well-defined” sets to which classical analysis had to be restricted. Thus the question was raised whether it is possible to extend the class of Borel sets to a wider class of sets with the same certainty. Lebesgue (*J. Math. Pures Appl.*, 1 (1905)) defined a function not belonging to any class of Baire functions by using the totality of †ordinals of the second class. (Later, this method was systematically developed as the †theory of sieves by Luzin.) However, it did not satisfy Borel as being effective. Can we, then, extend the Borel sets without any use of ordinals of the second class? The discovery of analytic sets gave an affirmative answer.

In this article (except in Section 1), we treat a space (denoted by X, Y, \dots) that is †homeomorphic to a †complete †separable †metric space and its subspace. Denote by \mathfrak{R} the **space of irrational numbers** (a metric space consisting of the irrational numbers $\in \mathbf{R}$ with the metric $|x - y|$ of x and y). The following properties of a subset S of a space X are equivalent: (i) S is a continuous image of \mathfrak{R} ; (ii) S is a continuous image of a Borel set in X ; (iii) S is the projection of a closed set in a product space $X \times \mathfrak{R}$; (iv) S is the projection of a Borel set in $X \times Y$. We call a set satisfying one of these properties an **analytic set**, an **A set**, or a Σ_1^1 set (in X). The complement of an analytic set is

called a **complementary analytic set** (or simply **coanalytic set**), a **CA set**, or a Π_1^1 set.

B. The Operation A and Sieves

When to each (n_1, \dots, n_k) of finite sequences of natural numbers there corresponds a unique element $E(n_1, \dots, n_k)$ of a family F of sets, this correspondence $\{E(n_1, \dots, n_k)\}$ is called a **schema of Suslin** (or **system of Suslin**) consisting of sets in F . Denoting an infinite sequence of natural numbers by $\{n_i\}$, the set given by $\bigcup_{i \in \mathbf{N}} \bigcap_k E(n_1, \dots, n_k)$ is called the **kernel** of a system of Suslin, and the operation of taking the kernel is called the **operation A** (**analytic operation**).

Let Q be the set of all rational numbers between 0 and 1 and F be a family of sets. Take a family $\{C_r\}_{r \in Q}$ of sets belonging to F with the index set Q (or more geometrically, a subset $C = \bigcup_{r \in Q} C_r \times r$ of $X \times Q$ when F is a family of subsets of a space X), and call it a **sieve** consisting of sets in F . Denoting by $\{r_i\}$ a (strictly) monotone decreasing sequence of elements of Q , we call the set $\bigcup_{i \in \mathbf{N}} \bigcap_k C_{r_k}$ (namely, the set of all x such that $C^{(x)} = \{r \mid (x, r) \in C\}$ is not well-ordered by the order \leq of rational numbers) the set obtained by a sieve C or the **sieved set** obtained by C . If the family F is closed with respect to countable intersection, then the family of all sets obtained by sieves consisting of sets in F is identical to the family of all sets obtained by applying the operation A to F . When F consists of all the closed sets in a given space, this is the family of all analytic sets. In particular, it is sufficient to take the family of closed intervals as F when X is the space of real numbers. Note that we can define sieves and sieved sets more generally by using the space of real numbers \mathbf{R} instead of the set Q of rationals.

C. Properties of Analytic Sets

It is evident from the definition that every †Borel set is analytic. If a Borel set is uncountable, then it is the union of a countable set and a one-to-one continuous image of \mathfrak{R} . The analyticity of sets is invariant under countable unions, intersections, and Cartesian products and the operation A and †Borel-measurable transformations. An uncountable analytic set contains a †perfect subset (Suslin). Therefore the possible cardinality of an analytic set is at most countable or is that of the continuum. Every analytic set enjoys the †Baire property, and in Euclidean space every analytic set is †Lebesgue measurable (Luzin, Sierpinski). If a set E in the Euclidean plane is analytic (coana-

lytic), then $I(E)$ is also analytic (coanalytic), where $I(E)$ is the set of all x such that the section $E^{(x)}$ of E that is parallel to the y -axis has a positive measure (M. Kondô and T. Tugué). Concerning the Baire property, the similar result for a set E of $X \times Y$ is obtained by replacing "to have a positive measure" by "to be the †second category (nonmeager) set" (A. S. Kechris). The Lebesgue measure of an analytic set is effectively calculable (in the sense that the measure of a Σ_1^1 (or $\Sigma_1^1(x)$) set is a Σ_1^1 (resp. $\Sigma_1^1(x)$) real number (- Section G and 356 Recursive Functions) (H. Tanaka). Every analytic (coanalytic) set E can be decomposed into \aleph_1 Borel sets. This decomposition is called a decomposition of E into constituents. An analytic (coanalytic) set is a Borel set if and only if it is decomposable into a countable number of constituents (Luzin, Sierpinski). In a space with the cardinality of the continuum, there exist analytic sets that are not Borelian. For example, in the space $C([0, 1])$ of continuous functions on the interval $[0, 1]$ (\rightarrow 168 Function Spaces) the set of all differentiable functions is coanalytic but not Borelian (S. Mazurkiewicz).

The following theorems are especially important in analytic set theory. **Luzin's first principle** (the first separation theorem): For every pair of disjoint analytic sets A, A_1 , there exists a Borel set B such that $A_1 \subset B$ and $B \cap A = \emptyset$. An immediate corollary of Luzin's first principle is **Suslin's theorem**: If both A and $X - A$ are analytic, then A is a Borel set. **Luzin's second principle** (the second separation theorem): For every pair of analytic sets A and B , there exist complementary analytic sets C and D such that $A - B \subset C$, $B - A \subset D$, and $C \cap D = \emptyset$. A one-to-one continuous image of a Borel set is Borelian (Suslin). More generally, for a given B -measurable function f defined on a Borel set B , the set A ($\subset f(B)$) of all points y whose inverse images $f^{-1}(y)$ are singletons is a complementary analytic set (**Luzin's unicity theorem**). In this theorem, we can replace "a singleton" by "a † σ -compact set" (V. Ya. Arsenin, K. Kunugui). Therefore, if a set is the image of a Borel set by a continuous function such that the inverse image of each point is a σ -compact set, then it is a Borel set.

D. Generalization to Projective Sets

A **projective set of class n** is inductively defined as follows: (i) the Borel sets are the projective sets of class 0; (ii) the projective sets of class $2n + 1$ are the continuous images of the sets of class $2n$; (iii) the projective sets of class $2n$ are the complements of the sets of class $2n - 1$.

The projective sets of class 1 are exactly the analytic sets, and those of class 2 are the complementary analytic sets. The following are fundamental properties of projective sets.

Denote by L_n the family of the projective sets of class n . Then (1) $L_{2n} \subset L_{2n+k}$ and $L_{2n+1} \subset L_{2n+2+k}$ ($k = 1, 2, \dots$). (2) the property of being a set of class n is invariant under countable unions, intersections, and Cartesian products and homeomorphisms; (3) a continuous image of a projective set of class $2n + 1$ is of the same class; (4) the projection on X of a set of class $2n + 1$ in $X \times Y$ is a set of class $2n + 1$ in X ; (5) the family of the projective sets of class $2n + 1$ in a space X is the family of the projections of all sets of class $2n$ in $X \times X$ (or $X \times \mathfrak{N}$); (6) the kernel of a system of Suslin consisting of sets of class n is a projective set of the same class, where $n \neq 0, 2$.

We frequently call a projective set of class $2n - 1$ a P_n set or a Σ_n^1 set, and that of class $2n$ a C_n set or a Π_n^1 set. A B_n set or a Δ_n^1 set is a set that is both P_n and C_n . The respective families of these sets are also denoted by $P_n(\Sigma_n^1)$, $C_n(\Pi_n^1)$, and $B_n(\Delta_n^1)$, respectively. In general, for a family \mathfrak{F} of sets in a space X , we denote by $C\mathfrak{F}$ the family of the complements $X - E$ of all sets E in \mathfrak{F} . We write $\text{Sep}_I(\mathfrak{F})$ and $\text{Sep}_{II}(\mathfrak{F})$ for the propositions obtained by substituting "set in \mathfrak{F} ," "set in \mathfrak{F} and in $C\mathfrak{F}$," and "set of $C\mathfrak{F}$ " for "analytic set," "Borel set," and "coanalytic set," respectively, in Luzin's first principle and Luzin's second principle, respectively. Furthermore, we say that the reduction principle holds for \mathfrak{F} , and denote it by $\text{Red}(\mathfrak{F})$, when for any sets $A, B \in \mathfrak{F}$ there exist $A_1, B_1 \in \mathfrak{F}$ such that $A_1 \subset A$, $B_1 \subset B$, $A_1 \cup B_1 = A \cup B$, and $A_1 \cap B_1 = \emptyset$. C. Kuratowski introduced the latter principle $\text{Red}(\mathfrak{F})$ which implies $\text{Sep}_I(C\mathfrak{F})$ and $\text{Sep}_{II}(C\mathfrak{F})$, and proved $\text{Red}(\Pi_1^1)$. Classically, reduction or separation principles for the projective sets of higher classes were not settled except for $\text{Sep}_{II}(\Pi_2^1)$, $\text{Sep}_{II}(\Pi_2^1)$ (P. S. Novikov), and $\text{Red}(\Sigma_2^1)$ (Kuratowski). Nowadays, it is known that these principles for $\{\Sigma_n^1, \Pi_n^1\}_{n \geq 3}$ are undecidable from the axiomatic set theory ZFC. If we assume †the axiom of constructibility $V = L$, then $\text{Red}(\Sigma_n^1)$ holds for $n \geq 3$ (J. W. Addison). On the other hand, under the assumption of †projective determinacy PD (\rightarrow Section H), it follows that $\text{Red}(\Pi_n^1)$ ($\text{Red}(\Sigma_n^1)$) holds when n is odd (even) (A. Martin, Addison and Y. N. Moschovakis).

E. Universal Sets

A set U in $\mathfrak{N} \times X$ is called the **universal set** for the projective sets of class n in X if for any projective set P of class n in X , there exists $z_0 \in \mathfrak{N}$ such that $P = \{x \mid (z_0, x) \in U\}$. Concern-

ing universal sets, we have the following result: for every $n > 0$, there exists a universal set for the projective sets of class n in X that is of the same class in $\mathfrak{N} \times X$. Hence in a space with the cardinality of the continuum, there exists a projective set of class $n + 1$ which is not of class n . In general, if a class \mathfrak{F} of sets is closed under taking a universal set and continuous preimages, then \mathfrak{F} and $C\mathfrak{F}$ cannot both have the reduction property. Therefore $\text{Red}(\Pi_n^1)$ or $\text{Red}(\Sigma_n^1)$ fails for each $n \geq 1$.

Any two universal sets of analytic sets are ${}^t\text{Borel}$ isomorphic. On the other hand, an analytic set E is not Borel isomorphic to any universal analytic set if the complement of E is an uncountable set without a perfect subset (A. Maitra and C. Ryll-Nardzewski). Hence, there are at least two equivalence classes of uncountable analytic sets with respect to the Borel isomorphism \cong under the assumption of $V=L$, since the latter assumption implies the existence of such a set E (- Section H).

F. Uniformization Principle

The uniformization problem arose during investigations of implicit functions. For a set E in a space $X \times Y$, **uniformization** of E is the finding of a subset V of E such that

$$\forall x(\exists y((x, y) \in E) \Leftrightarrow \exists! y((x, y) \in V)),$$

where $\exists! y$ is the t quantifier which means “there exists exactly one y .” A Borel set can be uniformized by choosing a suitable coanalytic set (Luzin). **Kondô’s uniformization theorem** (*Japan. J. Math.*, 15(1938)) is the most important result in descriptive set theory: every coanalytic set is uniformizable by a coanalytic set. As a corollary to this, any Σ_2^1 set is uniformizable by a Σ_2^1 set, and every Σ_2^1 set is obtained as a one-to-one continuous image of a coanalytic set.

Kondô’s original proof was very difficult to understand. Since t effective descriptive set theory was introduced by Kleene and Addison, the proof has been simplified by Addison, and a more elegant one has been given by J. R. Shoenfield [7, p. 188]. Nowadays, the theorem is also called the Novikov-Kondô-Addison theorem, and is one of the most powerful and fundamental theorems not only in descriptive set theory, but also in the t foundations of mathematics.

The uniformization of an analytic set is, in general, not to be found among analytic or coanalytic sets. There was a conjecture that any analytic set is uniformizable by specifying an A , set (difference of two analytic sets). Recently, this conjecture was negatively settled (J. Steel, Martin). Assuming that $V=L$, the

uniformization of a Σ_n^1 ($n \geq 3$) set is determined by specifying a Σ_n^1 set, and that of a Π_n^1 ($n \geq 2$) set by specifying a $\Pi_{n,p}^1$ set. On the other hand, if an axiom system of set theory (e.g., ZF; \rightarrow 33 Axiomatic Set Theory) is consistent, then it is still consistent even if we add to it the following proposition: There exists a Π_2^1 set whose uniformization is impossible by any choice of a definable set in the system (P. J. Cohen, A. Lévy). However, if we assume the existence of a t measurable cardinal (MC), it is known that every Π_2^1 set is uniformizable by a Π_2^1 set (Martin and R. M. Solovay). On the other hand, PD implies “every $\Pi_n^1(\Sigma_n^1)$ set is uniformizable by a $\Pi_n^1(\Sigma_n^1)$ set for each odd (resp. even) n ” (Moschovakis).

There are sufficient conditions on sections $E^{(x)}$ of a Borel set E in the Cartesian product $X \times Y$ for E to be uniformizable by the Borel set as follows: all sections $E^{(x)}$ are (i) countable (Luzin, Novikov), (ii) a -compact (Arsenin, Kunugui); or, as “large section property,” (iii) in the second category (S. K. Thomason, P. G. Hinman), or (iv) of positive measure (Tanaka, G. E. Sacks). For applications of descriptive set theory to analysis, an important uniformization result is **von Neumann’s selection theorem**: for any A set $E \subset X \times Y$, there exists a ${}^t\text{Baire}$ measurable and t absolutely measurable function $f: X \rightarrow Y$ (- 270 Measure Theory L (vi)) such that

$$\forall x(\exists y((x, y) \in E) \Leftrightarrow (x, f(x)) \in E).$$

Concerning implicit functions, any Borel set $E \subset X \times Y$ such that all sections $E^{(x)}$ are at most countable is expressed by a union of intersections of E with graphs of some Baire functions $f_n: X \rightarrow Y$, $n \in \mathbb{N}$ (Luzin). Recently, it was shown known as a generalization of this that any Borel set $E \subset X \times Y$ such that all sections are a -compact is a union of countably many Borel sets E_n for each of which all sections $E_n^{(x)}$ are compact (J. Saint-Raymond).

G. Kleene’s Hierarchy and Effectiveness

First, projective set theory in any space is reducible to the theory in the space of irrational numbers. Second, if we introduce a t weak topology in the set $\mathbb{N}^{\mathbb{N}} = {}^{\omega}\omega$ of t number-theoretic functions α with one argument, the resulting topological space $\mathbb{N}^{\mathbb{N}}$ is homeomorphic to the ${}^t\text{Baire}$ zero-dimensional space of irrational numbers. Third, any subset B of $\mathbb{N}^{\mathbb{N}}$ is open and closed in this topology if and only if there exist a function $\xi \in \mathbb{N}^{\mathbb{N}}$ and a predicate $A^{\xi}(\alpha)$ that is t general recursive in ξ such that $\alpha \in B \Leftrightarrow A^{\xi}(\alpha)$. Fourth, logical operations such as \neg , \vee , \wedge , $\exists x$ (where x is a variable ranging over the natural numbers), and $\exists \alpha$

exactly correspond to the operations (on sets) complementation, union, intersection, countable union, and projection, respectively. On the basis of these facts, projective set theory is regarded as the theory of the $\mathbb{N}^{\mathbb{N}}$ analytic hierarchy of Kleene. Here, the following example is remarkable: We can construct a Σ_1^1 set which is universal for the analytic sets (namely, a $\Sigma_1^1[\mathbb{N}^{\mathbb{N}}]$ set) in the space of irrational numbers (\rightarrow 356 Recursive Functions H).

The connection between projective set theory and logic has been discussed by C. Kuratowski and A. Tarski. From their point of view, semi-intuitionists such as Borel regard the set of natural numbers to be precisely clear in itself and also the continuum to be immediately recognizable by our geometric intuition. In their argument rational numbers do not play such an important role. They take, a priori, the set of irrational numbers as the fundamental domain, and intervals with rational extremities as the simplest sets of points among the subsets of the fundamental domain as the starting point of their argument. Here, the fundamental domain or each interval is not conceived as a totality of its elements, but recognized as a "uniform extent." In contrast to this, singletons and individual irrational numbers are not so simple. For this reason Borel introduced the notion of calculable numbers to study definable real numbers. Following Luzin, we say that a **calculable number** is a constructible real number in the sense that we can give it by an arithmetical approximation as precisely as we want. Now, this notion is nearly identical to the notion of an effectively calculable real number given by A. Church or A. M. Turing.

In the mathematics of the semi-intuitionists, the word "effective" has played an especially important role. Although these mathematicians have always agreed not to accept the axiom of choice, the exact meaning of "effective" has differed slightly among different members of their group or in different stages of the development of the theory. Such differences mainly arose in connection with the question: How can we tell whether given entities are finitary or individual? One way to guess the original intention held by Borel and others when they used the term "effective" is to replace the term by "recursive." Nowadays, the concept of "effectiveness" is used in this sense (\rightarrow 356 Recursive Functions C), and all classical results in descriptive set theory essentially have effective versions (or refinements via relativization). For example, the Novikov-Kondô-Addison theorem is described as follows: Any $\Pi_1^1(\Pi_1^1(\alpha))$ for an $\alpha \in {}^\omega\omega$ set is uniformizable by a Π_1^1 (resp. $\Pi_1^1(\alpha)$) set. Similarly, Suslin's theorem is that every analytic and

coanalytic (i.e., $\Delta_1^1(\alpha)$) set A is a Borel set of class $v(x)$ which is effective relative to the "definability α " of given A , where $v(a)$ is the +Constructive ordinal relative to α (Kleene, Tugué and Tanaka, A. Louveau, etc.).

H. Further Results in Axiomatic Set Theory and Strong Axioms

The recent development of axiomatic set theory has yielded the following propositions. Under the assumption that $V=L$ (K. Gödel, Novikov, Addison): (1) there exists an uncountable Π_1^1 set that does not contain any perfect subset; (2) there exists a nonmeasurable Δ_2^1 set; (3) there exists a Δ_2^1 set that does not have the Baire property. On the other hand, if the axioms of set theory ZFC plus "there exists an inaccessible cardinal number" are consistent, then so is "every projective set is Lebesgue measurable, has the Baire property, and has a perfect subset when it is uncountable" with them (Solovay). Concerning these properties, Martin's axiom MA and " $2^{\aleph_0} > \aleph_1$ " ($\text{MA} + \neg \text{CH}$ is consistent with ZFC; \rightarrow 33 Axiomatic Set Theory E) implies that every Σ_2^1 is measurable, and has the Baire property (Martin and Solovay). The possible cardinality of a Σ_2^1 set is at most \aleph_1 or that of the continuum, which is implied from "every Σ_2^1 set is a union of \aleph_1 Borel sets" (Sierpinski). It is known that one cannot prove or refute from ZFC the converse of the latter statement. If one assumes MC (the existence of a measurable cardinal), then every Σ_3^1 set is a union of \aleph_2 Borel sets. Therefore MC implies that the possible cardinality of a Σ_3^1 set is at most \aleph_2 or that of the continuum.

With each set $A \subset {}^\omega\omega$ we associate the following infinite game of perfect information $G(A)$, played by two players I and II. First player I chooses $n_0 \in \omega$, then player II chooses $n_1 \in \omega$, then I chooses $n_2 \in \omega$, and so on. The game ends after ω steps. Let $r(k) = n_k$; if $x \in A$, then I wins $G(A)$; otherwise II wins. A strategy (for I or II) is a rule that tells the player what move to make, depending on the previous moves of both players. A winning strategy is one such that the player who follows it always wins. The game $G(A)$ is determined if one of the players has a winning strategy. The axiom of **projective determinacy** PD is the assertion that for every projective set $A (\subset {}^\omega\omega)$ the game $G(A)$ is determined. PD is widely used to solve problems on projective sets that are not decidable from ZFC, and is a plausible hypothesis that does not seem to contradict the axiom of choice at present; and it has pleasing consequences in descriptive set theory. For $n \geq 3$,

reductions and uniformization results of Σ_n^1 (Π_n^1) sets mentioned earlier are such examples. Furthermore, PD implies that every projective set is Lebesgue measurable, has the Baire property, and contains a perfect subset if it is uncountable.

Now, write Determinacy (Σ_n^1) (Determinacy (Δ_n^1)), or simply, $\text{Det}(\Sigma_n^1)$ ($\text{Det}(\Delta_n^1)$), when $G(A)$ is determined for every Σ_n^1 (Δ_n^1) set A . Using the fact that the open game is determined, namely $\text{Det}(\Sigma_1^0)$ (D. Gale and F. M. Stewart), D. Blackwell has given a short proof of $\text{Sep}_1(\Sigma_1^1)$. The extremely difficult problem was to show $\text{Det}(\Delta_1^1)$ in ZF (indeed, one cannot prove $\text{Det}(\Delta_1^1)$ only from the Zermelo axioms $z; \rightarrow 33$ Axiomatic Set Theory B). In 1975, Martin solved this problem. Concerning Σ_1^1 , $\text{Det}(\Sigma_1^1)$ cannot be proved from only the ZF axioms, but MC implies $\text{Det}(\Sigma_1^1)$ (Martin). However, it is known that $\text{Det}(\Sigma_2^1)$ cannot be proved even if we assume MC.

By a pre-well-ordering, we mean a relation \leq having all the properties of a well-ordering except for \dagger antisymmetry. Denote by δ_n^1 the least ordinal that cannot be the length of a prewell-ordering on ${}^\omega\omega$ belonging to Δ_n^1 . It is a classical result that $\delta_1^1 = \omega_1 (= K)$. The following results are mainly due to Martin: $\delta_2^1 \leq \aleph_2$, MC $\Rightarrow \delta_3^1 \leq \aleph_3$, and also, PD $\Rightarrow \delta_4^1 \leq \aleph_4$. Let κ be an infinite cardinal. A set $A \subset X$ is called κ -Suslin when there exists a closed set $C \subset X \times {}^\omega\kappa$ (${}^\omega\kappa = \{f \mid f: \omega \rightarrow \kappa\}$) such that

$$x \in A \Leftrightarrow \exists f [f \in {}^\omega\kappa \wedge (x, f) \in C].$$

A is \aleph_0 -Suslin iff $A \in \Sigma_1^1$. If A is a Σ_2^2 set, then A is \aleph_1 -Suslin (Shoenfield). Under the assumptions of MC, if A is a Σ_3^1 set, then it is \aleph_2 -Suslin (Martin). Furthermore, $\text{Det}(\Delta_{2n}^1)$ implies the following facts (Kechris, Moschovakis): every Σ_{2n+2}^1 set is κ -Suslin, where κ is the cardinal of δ_{2n+1}^1 ; every Σ_{2n+1}^1 set is κ -Suslin for some $\kappa < \delta_{2n+1}^1$.

The axiom of determinacy AD is the assertion that the game $G(A)$ is determined for every set A . AD implies that every set is measurable (J. Mycielski and S. Swierczkowski), has the Baire property, and has a perfect subset if it is uncountable (Morton Davis). Though AD contradicts the axiom of choice, some consequences of AD in the area of projective sets are more desirable than the consequences of the axiom of choice. The following are examples of consequences of AD. (1) A is a Σ_2^1 set iff A is a union of \aleph_1 Borel sets (Moschovakis). (2) For each n , δ_n^1 is a cardinal (Moschovakis). (3) Generalization of Suslin's theorem holds as follows: $\mathcal{A}_n^1 = \mathbb{B}_{\delta_n^1}$ (= the smallest \dagger Boolean algebra containing the open sets and closed under complementation and unions of length $< \delta_n^1$) (Martin, Moschovakis).

I. Polish Spaces, Luzin Spaces, and Suslin Spaces

A topological space homeomorphic to a complete separable metric space is called a **Polish space**. A subspace E of a Polish space X is Polish if and only if it is a G_δ -subset of X , i.e., a countable intersection of open subsets of X (Aleksandrov and Uryson). A Hausdorff topological space X is called a **Luzin space** (resp. **Suslin space**) if we can find a Polish space S and a continuous bijective (resp. surjective) mapping $f: S \rightarrow X$. Every Polish space is a Luzin space and every Luzin space is a Suslin space.

Let X be a Hausdorff topological space. A subset E of X is called a **standard set** (resp. **analytic set**) if the set E with the relative topology is a Luzin (resp. Suslin) space. The analytic subsets are closed under analytic operations. Every analytic set is obtained from closed sets by applying the analytic operation. Every analytic set is \dagger universally measurable. Every standard set is a Borel set. Every countable union or intersection of standard sets is standard.

A subset E of a Suslin space is a Borel set if and only if both E and E^c are analytic. A subset E of a Luzin space is a Borel set if and only if E is standard.

Let f be a Borel measurable mapping from a Suslin space X into another Suslin space Y . Then the image $f(X)$ is an analytic subset of Y . Furthermore, if f is injective, then f^* gives a \dagger Borel isomorphism between X and $f(X)$. If both X and Y are Luzin spaces and if f is injective, then $f(X)$ is a Borel subset of Y . Every Suslin space is \dagger Borel isomorphic to an analytic subset of \mathbf{R} , and every Luzin space is Borel isomorphic to one of the following spaces: (1) $\mathbf{N}_n = \{1, 2, \dots, n\}$, (2) $\mathbf{N} = \{1, 2, \dots\}$, (3) \mathbf{R} .

The selection theorem due to von Neumann (- Section F) holds when X and Y are Suslin spaces. This fact and its ramifications [15, 16] are useful in nonlinear functional analysis and control theory.

Practically all useful spaces appearing in functional analysis and probability theory are Polish or Luzin spaces. Examples: (i) Every locally compact Hausdorff space with a countable open base is Polish. (ii) Every separable Banach space is Polish. (iii) The set C of all continuous functions on $[0, 1]$ with the \dagger topology of uniform convergence is a Polish space. (iv) The set D of all right continuous functions on $[0, 1]$ with left limits is a Polish space when it is endowed with the Skorokhod topology (- 250 Limit Theorems in Probability Theory E). (v) The space \mathcal{D}' of distributions and the space \mathcal{S}' of tempered distri-

butions (- 125 Distributions and Hyperfunctions) are Luzin spaces.

References

- [1] C. Kuratowski, Topologie I, Warsaw, revised edition, 1948.
- [2] A. A. Lyapunov (Liapunov), E. A. Shchegol'kov (Stschegolkow), and V. J. Arsenin, Arbeiten zur deskriptiven Mengenlehre, Deutscher Verlag der Wissenschaften, 1955.
- [3] N. N. Luzin, Sur les ensembles analytiques, Fund. Math., 10 (1927), 1-95.
- [4] N. N. Luzin, Leçons sur les ensembles analytiques et leurs applications, Gauthier-Villars, 1930.
- [5] W. Sierpiński, Les ensembles projectifs et analytiques, Mémor. Sci. Math., Gauthier-Villars, 1950.
- [6] H. Rogers, Jr., Theory of recursive functions and effective computability, McGraw-Hill, 1967, ch. 16.
- [7] J. R. Shoenfield, Mathematical logic, Addison-Wesley, 1967, ch. 7.
- [S] D. A. Martin, Descriptive set theory: Projective sets, Handbook of Mathematical Logic, North-Holland, 1977, ch. C.8.
- [9] T. Jech, Set theory, Academic Press, 1978, pt. IV.
- [10] P. G. Hinman, Recursion-theoretic hierarchies, Springer, 1978, pt. B.
- [11] Y. N. Moschovakis, Descriptive set theory, North-Holland, 1980.
- [12] C. A. Rogers, J. E. Jayne, C. Dellacherie, F. Topsøe, J. Hoffman-Jørgensen, D. A. Martin, A. S. Kechris, and A. H. Stone, Analytic sets, Academic Press, 1980.
- [13] N. Bourbaki, Éléments de mathématique, Topologie générale, Hermann, 1974, ch. X, sec. 6.
- [14] J. Hoffmann-Jørgensen, The theory of analytic spaces, Various publication series 10, Math. Institute, Aarhus Univ., 1970.
- [15] M.-F. Sainte-Beuve, On the extension of von Neumann-Aumann's theorem, J. Functional Anal., 17 (1974), 112-129.
- [16] A. F. Filippov, On certain questions in the theory of optimal control, SIAM J. Control, 1 (1962), 76-84.

23 (X1.21) Analytic Spaces

A. General Remarks

An \dagger analytic function of a complex variable has as its natural domain of definition a \dagger Rie-

mann surface, i.e., a 1-dimensional complex analytic manifold. In the case of several complex variables, the set of zeros of an analytic function, the quotient space of a domain by a \dagger properly discontinuous group of analytic automorphisms, the existence domain of an \dagger algebroidal function, etc., are, strictly speaking, not necessarily complex analytic manifolds. It is necessary to consider a more general category of complex analytic manifolds with singularities, and the notion of analytic spaces is drawn from these examples. Many of the properties of complex analytic manifolds are extended to analytic spaces; on the other hand, theories specific to analytic spaces have also been developed.

B. Analytic Sets

We say that a subset A of a complex analytic manifold G is an **analytic set** in G if it is a closed subset and each point of A has a neighborhood U such that $U \cap A$ is the set of common zeros of a finite number of holomorphic functions in U . Specifically, if A is locally the set of zeros of a single holomorphic function that does not vanish identically, then A is called **principal**. Two subsets S_1 and S_2 of G are called equivalent at $z^0 \in G$ if there exists a neighborhood U of z^0 such that $S_1 \cap U = S_2 \cap U$. By this equivalence relation, every subsets S of G defines its germ S_{z^0} at z^0 . A **germ of an analytic set** at z^0 is the germ at z^0 of an analytic set in a neighborhood of z^0 . Each germ A , of an analytic set at $z^0 = 0 \in G$ is associated with an ideal $I(A_0) = \{f \mid f \in H(0), f|_A = 0\}$ in the ring $H(0)$ of \dagger germs of holomorphic functions at 0 . We call A , **reducible** if A is the union of two germs of analytic sets A_1 and A_2 with $A_0' \neq A_0$, $A_0'' \neq A_0$; otherwise, A is called **irreducible**. An analytic set A is called **irreducible at 0** if the germ at 0 of A is irreducible. Properties of A , and $I(A_0)$ correspond to each other. Thus A is irreducible if and only if $I(A_0)$ is prime.

As the ring $H(0)$ is \dagger Noetherian, in a neighborhood of every point z^0 an analytic set A is represented as a union of a finite number of analytic sets A_i that are irreducible at z^0 . These A_i are essentially unique. If an analytic set A is irreducible at z^0 , then there exists a system of local coordinates (z_1, \dots, z_n) centered at z^0 and a pair of natural numbers $d \leq n$ and k such that, in a neighborhood of z^0 , A is a k -sheeted **ramified covering space** with covering mapping $\varphi: (z_1, \dots, z_n) \rightarrow (z_1, \dots, z_d)$; i.e., for an analytic set R in a neighborhood of $0 \in \mathbb{C}^d$, $\varphi: A - \varphi^{-1}(R) \rightarrow \mathbb{C}^d - R$ is, in a neighborhood of z^0 , a k -sheeted covering mapping, where $A - \varphi^{-1}(R)$ is a connected d -dimensional com-

plex analytic manifold in the neighborhood. The coordinates of the k points of $A - \varphi^{-1}(R)$ over each point (z_1, \dots, z_d) are holomorphic in these variables. The number d is the **(local) dimension** of A at z^0 and is denoted by $\dim_{z^0} A$. From this local representation, we obtain **Rückert's zero-point theorem**: Every prime ideal \mathfrak{P} in $H(0)$ is equal to $I(A_0)$, with A , an irreducible germ of an analytic set at 0. In this case the dimension at 0 of an analytic set A that defines A , is equal to the \dagger Krull dimension of the \dagger local ring $H(0)/\mathfrak{P}$. The theory of local rings is very important in the study of germs of analytic sets. The dimension of a general analytic set A at z^0 is defined by $\dim_{z^0} A = \sup_i \dim_{z^0} A_i$, where $A = \bigcup_i A_i$ in a neighborhood of z^0 , with the A_i irreducible at z^0 . If $\dim_{z^0} A_i$ is equal to d for all i , then A is called **purely d -dimensional** at z^0 . The **(global) dimension** of A is defined by $\dim A = \sup_{z \in A} \dim_{z^0} A$. A **purely d -dimensional analytic set** is defined to be an analytic set that is purely d -dimensional at every one of its points.

A point z^0 of an analytic set A is called **ordinary (regular or simple)** if A has the structure of a complex analytic submanifold in a neighborhood of z^0 . The set A^* of ordinary points of A is dense and open in A . The set $A^* = A - A^*$ of **singular** (not ordinary) points is an analytic set in G . If A is purely d -dimensional, then A^* is a d -dimensional complex analytic manifold and A^* is an analytic set of dimension $\leq d - 1$.

Let A be an analytic set of dimension d in G , and B a purely d' -dimensional analytic set in $G - A$ with $d' > d$. Then the closure \bar{B} of B in G is a purely d' -dimensional analytic set (**Remmert-Stein continuation theorem** [17]).

For every analytic set A in G , the \dagger analytic sheaf $\mathcal{S}(A)$ of germs of holomorphic functions over G that vanish on A is \dagger coherent (H. Cartan). We call $\mathcal{S}(A)$ the sheaf of ideals defined by an analytic set A . Let \mathcal{O}_G be the sheaf of germs of holomorphic functions over G . Then $\mathcal{O}_A = (\mathcal{O}_G / \mathcal{S}(A))|_A$ is a coherent sheaf of rings over A . \mathcal{O}_A is called the sheaf of germs of holomorphic functions on an analytic set A .

C. Analytic Spaces

A \dagger ringed space $(X, 0_*)$ with Hausdorff base space X is called an **analytic space** if for every point $x \in X$, there exists an open neighborhood U of x such that the ringed space $(U, \mathcal{O}_X|_U)$ is isomorphic to a ringed space (A, \mathcal{O}_A) , where A is an analytic set in an open set G of some \mathbb{C}^n . The structure sheaf \mathcal{O}_X is then called a **sheaf of germs of holomorphic functions**. The notion of \dagger holomorphic mapping from one open set in \mathbb{C}^n into another is generalized to the case of

mappings from one analytic set into another. An analytic set Y in an analytic space X and the sheaf \mathcal{O}_Y of germs of holomorphic functions on Y are defined as in the case where X is a complex manifold. The ringed space $(Y, 0_*)$ is an analytic space and is called an **analytic subspace** of X . For an analytic space X , the notions of $\dim X$, $\dim X$, irreducibility, and pure dimensionality are defined as for an analytic set $A \subset G \subset \mathbb{C}^n$. Every analytic space X is the union of a locally finite family of irreducible analytic subspaces X_i , called the **irreducible components** of X .

Let $\varphi: X \rightarrow Y$ be a holomorphic mapping of an analytic space X into another, Y . Its **rank** at $x \in X$ is defined by $r_\varphi(x) = \dim_x X - \dim_x \varphi^{-1}(\varphi(x))$. The number $r_\varphi = \sup_{x \in X} r_\varphi(x)$ is called the **rank** of φ . The **set of degeneracy** E_φ of φ is the set of points $x \in X$ such that $r_{\varphi|_{X'}}(x) < r_{\varphi|_{X'}}$ for an irreducible component X' of X through x . The mapping φ is **non-degenerate** if $E_\varphi = \emptyset$. For any $k \in \mathbb{N}$, $\{x \in X \mid r_\varphi(x) \leq k\}$ is an analytic set (R. Remmert [16]). In particular, E_φ is analytic. For a holomorphic mapping $\varphi: X \rightarrow Y$, the inverse image of an analytic set in Y is an analytic set in X . However, the image of an analytic set is not necessarily analytic. If φ is \dagger proper, then the image $\varphi(X')$ of an analytic set X' in X is an analytic set in Y of dimension $r_{\varphi|_{X'}}$ and is irreducible if X' is irreducible (**Remmert's theorem** [16]).

D. Modifications and Resolution of Singularities

Let M be a subset of an analytic space X . If, for every point $x \in X$, there exists an open neighborhood U of x and an analytic set M^* in U , containing $U \cap M$ such that $U - M^*$ is dense in U , then M is called **analytically thin**. Let $\varphi: X \rightarrow Y$ be a holomorphic mapping. Suppose that there exist two analytically thin sets $A \subset X$, $N \subset Y$ such that φ induces an isomorphism between $X - M$ and $Y - N$. Then X is called a **holomorphic modification** of Y . If furthermore φ is proper, then X is called a **proper modification** of Y . A **monoidal transformation** of an analytic space X with respect to a coherent sheaf of ideals \mathcal{I} is defined as in the case of a complex manifold or an algebraic variety (\rightarrow 16 Algebraic Varieties L; 72 Complex Manifolds H). It is a proper modification $f: X^* \rightarrow X$ such that the inverse image ideal sheaf $f^{-1}\mathcal{I} \cdot \mathcal{O}_{X^*}$ is invertible, and is universal among all proper holomorphic mappings $h: Z \rightarrow X$ with the property that $h^{-1}\mathcal{I} \cdot \mathcal{O}_{Z^*}$ is invertible, where $f^{-1}\mathcal{I}$ is the \dagger inverse image of \mathcal{I} . If \mathcal{I} is the sheaf of ideals defined by an analytic set Y in X , the mon-

oidal transformation of X with respect to \mathcal{J} is often called the **blowing-up** of X with center Y .

H. Hironaka [14] proved that, if X is an analytic space which is countable at infinity (i.e., a countable union of compact sets), then there is a proper modification $\pi: X' \rightarrow X$ with X' smooth (i.e., free from singular points). Such a modification is called a **desingularization** (or **resolution of singularities**) of X . Moreover, over any relatively compact open set U of X , π is the product of a finite sequence of blowing-ups $\pi_i: X'_i \rightarrow X'_{i-1}$ ($X'_0 = X$), with smooth centers Y'_{i-1} along which X'_{i-1} is \dagger normally flat. This deep result enables one to derive properties of analytic spaces from those of complex manifolds.

Let X and Y be two analytic spaces and G an analytic set in $X \times Y$. If the canonical projection $\pi: G \rightarrow X$ is a holomorphic (or proper) modification, then we say that a **meromorphic mapping** (a proper meromorphic mapping) μ of X into Y is defined. The set G is then called the **graph** of μ . A holomorphic mapping $\varphi: X \rightarrow Y$ can be viewed as a proper meromorphic mapping. Let $\mu: X \rightarrow Y$ be a proper meromorphic mapping. Then $\mu(x)$ (the projection of $\pi^{-1}(x)$ into Y) is a nonempty analytic set in Y for every point $x \in X$. Moreover, there exists an analytic set N , with $X \setminus N$ dense in X , such that μ maps $X \setminus N$ into Y holomorphically. The smallest set N with this property is called the **set of points of indeterminacy** or the **singularity set** of μ . A meromorphic mapping $f: X \rightarrow \mathbf{P}^1(\mathbf{C})$ is called a **meromorphic function** on X if none of the irreducible components of X is mapped to $\{\infty\}$ by f . The set $f^{-1}(0) = \pi((X \times \{0\}) \cap G)$ is called the set of zero points of X , and the set $f^{-1}(\infty)$ is called the set of poles. These are analytic sets in X . Let f_1, \dots, f_k be meromorphic functions on X . Then, by a suitable proper modification of X , one can eliminate the points of indeterminacy of the meromorphic mapping $f: X \rightarrow (\mathbf{P}^1(\mathbf{C}))^k$ defined by $x \mapsto (f_1(x), \dots, f_k(x))$, i.e., one can modify f to be holomorphic. The ring of meromorphic functions on X is invariant under proper modifications of X . If X is irreducible and compact, then the field of meromorphic functions on X is a simple algebraic extension of the \dagger field of rational functions of k ($\leq \dim X$) variables.

Let $(X, 0_*)$ be an analytic space. A point $x \in X$ is called **normal** for X if $\mathcal{O}_{X,x}$ is a \dagger normal local ring. The set of nonnormal points for X is an analytically thin analytic set in X (K. Oka). Every ordinary point of X is normal. We call X **normal** if every one of its points is normal. Every nondegenerate holomorphic mapping of an irreducible X into an irreducible and normal Y is an open mapping if its rank is equal to the dimension of Y (Remmert).

For every analytic space X , there exists a proper modification $v: \tilde{X} \rightarrow X$ with \tilde{X} normal such that v is nondegenerate and $v: \tilde{X} \rightarrow v^{-1}(S)$ is an isomorphism, where S is the set of non-normal points. Such a proper modification of X is unique up to isomorphisms. We call \tilde{X} a **normalization** of X with normalizing mapping v .

Let $\varphi: X \rightarrow Y$ be a holomorphic modification. Suppose that Y is normal at $\varphi(x^0)$ ($x^0 \in X$). If the set $\varphi^{-1}(\varphi(x^0))$ contains an isolated point, then $\varphi^{-1}(\varphi(x^0)) = x^0$, and φ is an isomorphism in a neighborhood of x^0 (an analog of \dagger Zariski's main theorem). In particular, if $\varphi: X \rightarrow Y$ is a holomorphic modification and Y is normal, then φ maps $X - E_\varphi$ isomorphically onto the dense open set $\varphi(X - E_\varphi)$. Furthermore, if a holomorphic mapping $\varphi: X \rightarrow Y$ is injective and X and Y are irreducible, normal, and n -dimensional, then $\varphi(X)$ is an open set in Y and $\varphi^{-1}: \varphi(X) \rightarrow X$ is holomorphic.

E. Analytic Spaces in the Sense of Behnke and Stein

Let $\varphi: \tilde{G} \rightarrow G$ be a proper continuous mapping of a connected \dagger locally compact space \tilde{G} onto a domain $G \subset \mathbf{C}^n$. The triple $\mathfrak{G} = (\tilde{G}, \varphi, G)$ is an **analytic covering space** over G if the following conditions are satisfied: (i) $\varphi^{-1}(z^0)$ is a finite set for every point $z^0 \in G$. (ii) There exists an analytic set $A \subset G$ of dimension $\leq n-1$ such that $\varphi: \tilde{G} - \varphi^{-1}(A) \rightarrow G - A$ is a local homeomorphism and every point of $\varphi^{-1}(A)$ has a fundamental system of neighborhoods U such that both U and $U - \varphi^{-1}(A) \neq \emptyset$ are connected. As \mathfrak{G} is unramified over $G - A$, the number of points in $\varphi^{-1}(z^0)$ is constant for $z^0 \in G - A$ and is called the **number of sheets** of \mathfrak{G} . A point $\tilde{z} \in \tilde{G}$ is called a **ramification point** of \mathfrak{G} if the restriction of φ to any neighborhood of \tilde{z} is not a homeomorphism. Denote by B the set of ramification points of \mathfrak{G} . Then $\varphi(B) \subset A$ is an analytic set of dimension $\leq n-1$. Let f be a continuous complex-valued function in an open set D in \tilde{G} . We call f **holomorphic** in D if for every point $\tilde{z}^0 \in D - B$ and for every open neighborhood V of $z^0 = \varphi(\tilde{z}^0)$ over which φ is a homeomorphism, $f \circ \varphi^{-1}$ is holomorphic in V . Denote by $\mathcal{O}_{\tilde{G}}$ the sheaf of germs of holomorphic functions over \tilde{G} . Then $(\tilde{G}, \mathcal{O}_{\tilde{G}})$ is a ringed space. An **analytic space in the sense of Behnke and Stein** is a Hausdorff ringed space $(X, 0_*)$ that is locally isomorphic to a ringed space of the form $(\tilde{G}, \mathcal{O}_{\tilde{G}})$ [3]. \dagger Riemann's theorem on removable singularities holds for such spaces. Every normal analytic space is an analytic space in the sense of Behnke and Stein. An analytic covering space $\mathfrak{G} = (\tilde{G}, \varphi, G)$

is a **C-covering space** (covering space in the sense of Cartan) if for every point $z^0 \in G$ there exist an open neighborhood V of z^0 and a holomorphic function g in $U = \varphi^{-1}(V)$ which can be defined by an irreducible polynomial of degree k such that its coefficients are holomorphic functions on V and the coefficient of its highest term is 1, where k is the number of sheets of \mathfrak{G} . A **C-analytic space** is an analytic space in the sense of Behnke and Stein that is locally isomorphic to a C-covering space. The category of C-analytic spaces coincides with that of normal analytic spaces. According to H. Grauert and Remmert [10], every analytic covering space is a C-analytic covering space. Therefore every analytic space in the sense of Behnke and Stein is a normal analytic space.

Let R be an equivalence relation in an analytic space X . Given a subset A of X , denote by $R[A]$ the set of points of X which are R -equivalent to points of A . We call R **proper** if $R[K]$ is compact for every compact set K in X . Let φ be a proper holomorphic mapping of an analytic space X into another Y . For $x, x' \in X$, let $x \equiv x'(R)$ be defined by $\varphi(x) = \varphi(x')$. The equivalence relation R is then proper. We consider the quotient space X/R and the canonical projection $p: X \rightarrow X/R$. With each open set U in the quotient space X/R we can associate the ring of holomorphic functions in $p^{-1}(U)$ that are constant on $p^{-1}(\bar{x})$ for every $\bar{x} \in U$. This leads to a ringed space $(X/R, \mathcal{O}_{X/R})$, which is proved to be an analytic space by **Grauert's theorem** [9]: All the \dagger direct images of a coherent analytic sheaf over X by a proper holomorphic mapping $\varphi: X \rightarrow Y$ are coherent. For every proper equivalence relation R in X , the ringed space $(X/R, \mathcal{O}_{X/R})$ is an analytic space if and only if for every point $\bar{x} \in X/R$ there exists an open neighborhood V of \bar{x} such that functions in $\Gamma(V, \mathcal{O}_{X/R})$ separate the points of V (H. Cartan).

F. Stein Spaces

For an analytic space (X, \mathcal{O}_X) let us consider the following conditions: (i) Functions in $I(X, \mathcal{O}_X)$ separate the points of X . (ii) X is **K-complete**; i.e., for every point $x \in X$ there exist a finite number of $f_i \in I(X, \mathcal{O}_X)$ ($i = 1, \dots, k$) such that the holomorphic mapping $f = (f_i): X \rightarrow \mathbb{C}^k$ is nondegenerate at x . (iii) Every compact analytic set in X is a finite set. Condition (i) implies (ii), and (ii) implies (iii). If an irreducible analytic space X is K-complete, then X is a countable union of compact sets (Grauert [S]). In fact, if $n = \dim X$, there exist functions $f_i \in \Gamma(X, \mathcal{O}_X)$ ($i = 1, \dots, n$) such that the holomorphic mapping $f = (f_i): X \rightarrow \mathbb{C}^n$ is nondegen-

erate. The notion of holomorphic convexity (\rightarrow 21 Analytic Functions of Several Complex Variables H) is carried over to analytic spaces. For a holomorphically convex analytic space, conditions (i), (ii), and (iii) are equivalent (Grauert [8]). A **Stein space** (or **holomorphically complete space**) is a holomorphically convex analytic space that satisfies one of the conditions (i), (ii), or (iii). In a holomorphically convex analytic space (X, \mathcal{O}_X) , let R be the equivalence relation defined by $I(X, \mathfrak{G})$; i.e., for $x, x' \in X$, $x \equiv x'(R)$ if and only if $f(x) = f(x')$ for every $f \in \Gamma(X, \mathcal{O}_X)$. Then R is proper. The analytic space $(X/R, \mathcal{O}_{X/R})$ is a Stein space. A Stein space is a generalization of the notion of a Stein manifold. Fundamental theorems A and B on Stein manifolds hold Verbatim for Stein spaces (\rightarrow 21 Analytic Functions of Several Complex Variables L). Therefore the main properties of Stein manifolds are inherited by Stein spaces. Let $\varphi: X \rightarrow Y$ be a holomorphic mapping of an analytic space X into another, Y . If for every $x \in X$ all the connected components of the fibers $\varphi^{-1}(\varphi(x))$ are compact, then the equivalence relation R' defined by those components (i.e., for x and x' in X , $x \equiv x'(R')$ if and only if x and x' belong to the same component of $\varphi^{-1}(\varphi(x))$) is proper, and the ringed space $(X/R', \mathcal{O}_{X/R'})$ is an analytic space. In particular, if X is a holomorphically convex irreducible analytic space and R is the equivalence relation defined by $\Gamma(X, \mathcal{O}_X)$, then all the fibers of the canonical projection $p: X \rightarrow X/R$ are connected.

G. Further Topics

The notion of analytic space can be generalized as follows (Grauert [9]). A ringed space (X, \mathcal{O}_X) is a **general analytic space** if it is locally isomorphic to a ringed space (A, \mathcal{H}_A) , where A is an analytic set in a domain $G \subseteq \mathbb{C}^n$, and $\mathcal{H}_A = (\mathcal{O}_G/\mathcal{I})|_A$ for some coherent analytic subsheaf \mathcal{I} of \mathcal{O}_G such that $\text{Supp}(\mathcal{O}_G/\mathcal{I}) (= \{z \in G \mid (\mathcal{O}_G/\mathcal{I})_z \neq 0\}) = A$. An **analytic subspace** of (X, \mathcal{O}_X) is a ringed space (Y, \mathcal{O}_Y) where $Y = \text{Supp}(\mathcal{O}_X/\mathcal{I})$ and $\mathcal{O}_Y = \mathcal{O}_X/\mathcal{I}$ for some coherent sheaf of ideals \mathcal{I} of \mathcal{O}_X . (Y, \mathcal{O}_Y) is also a general analytic space. A. Douady [6] showed that for any general analytic space (X, \mathcal{O}_X) there exists a natural structure of a general analytic space on the totality of all the compact analytic subspaces of (X, \mathcal{O}_X) . The resulting analytic space is called the **Douady space** of (X, \mathcal{O}_X) . For the proof, the notion of a **Banach analytic space** is used, which is obtained by first defining analytic subspaces in an open subset of a complex \dagger Banach space and then patching them. A Douady space can

be used, for example, to show that the identity component of the analytic automorphism group of a compact Kähler manifold is a \dagger complex Lie group which is naturally an extension of a \dagger complex torus by a \dagger linear algebraic group (A. Fujiki, D. Lieberman; \rightarrow 232 Kähler Manifolds C).

References

- [1] S. S. Abhyankar, Local analytic geometry, Academic Press, 1964.
 - [2] C. Banica and O. Stanasila, Algebraic methods in the global theory of complex spaces, Wiley, 1976.
 - [3] H. Behnke and K. Stein, Modifikation komplexer Mannigfaltigkeiten und Riemannscher Gebiete, Math. Ann., 124 (1951), 1-16.
 - [4] L. Bers, Introduction to several complex variables, Courant Institute, 1963.
 - [5] Sémin. H. Cartan, Ecole Norm. Sup., 1953-1954, 1957-1958, 1960-1961.
 - [6] A. Douady, Le problème de modules pour les sous-espaces analytiques complexes d'un espace analytique donné, Ann. Inst. Fourier, 16 (1966), 1-95.
 - [7] G. Fischer, Complex analytic geometry, Springer, 1976.
 - [8] H. Grauert, Charakterisierung der holomorph-vollständiger komplexen Räume, Math. Ann., 129 (1955), 233-259.
 - [9] H. Grauert, Ein Theorem der analytischen Garbentheorie und die Modulräume komplexer Strukturen, Publ. Math. Inst. HES, 5 (1960).
 - [10] H. Grauert and R. Remmert, Komplexe Räume, Math. Ann., 136 (1958), 245-318.
 - [11] H. Grauert and R. Remmert, Analytische Stellenalgebren, Springer, 1971.
 - [12] R. C. Gunning and H. Rossi, Analytic functions of several complex variables, Prentice-Hall, 1965.
 - [13] M. Hervé, Several complex variables, local theory, Tata Inst. Studies in Math., Oxford Univ. Press, 1963.
 - [14] H. Hironaka, Desingularization of complex-analytic varieties, Act. Congr. Intern. Math., Nice, 1972, Gauthier-Villars, vol. 2, 6277632.
 - [15] R. Narashimhan, Introduction to the theory of analytic spaces, Lecture notes in math. 25, Springer, 1966.
 - [16] R. Remmert, Holomorphe und meromorphe Abbildungen komplexer Räume, Math. Ann., 133 (1957), 3288370.
 - [17] R. Remmert and K. Stein, Über die wesentlichen Singularitäten analytischer Mengen, Math. Ann., 126 (1953), 2633306.
- Also \rightarrow references to 21 Analytic Functions of Several Complex Variables.

24 (Xx1.1) Ancient Mathematics

A. General Remarks

To determine the beginning of the history of mathematics, one must define the term "mathematics." Only speculation based on the observation of primitive peoples today can be made regarding the development of the number concept among prehistoric peoples. The prehistoric period ended in Egypt and Mesopotamia c. 3000 B.C., and a little later in the valleys of the large rivers in India and China.

Since the basis of the civilizations in the river valleys of the ancient world was agriculture, the administrators first had to control watering systems through irrigation, drainage, pumping, and canalization; second, they had to measure land and harvests for tax collection; and third, they had to establish a calendar by observation of the heavenly bodies. All these tasks demanded some knowledge of mathematics.

Additional knowledge of mathematics was certainly needed for construction of the palaces and tombs. We know something of the development of mathematical knowledge during these ages from some recovered artifacts, but there remains the possibility of new finds that will bring about a basic change in our knowledge of the history of mathematics during this period.

B. Mathematics in Egypt

The main sources for our understanding of the history of mathematics in Egypt are the Moscow papyrus and the more important Rhind papyrus, both discovered in the 19th Century. The Greeks place the origin of their mathematics in Egypt, but it seems that Egyptian mathematics was limited to practical mathematics. The Egyptians had a decimal numeration system, but the place value was not clear; they used fractions, which they always decomposed into the sums of unit fractions (i.e., fractions with 1 as numerator); they solved the problems of everyday arithmetic that were reducible to linear equations; they computed approximate areas and volumes of some figures for the purpose of measurement of farmland or granaries and for construction work; they had exact formulas for the computation of areas of triangles and of trapezoids; and they used $(16/9)^2 = 3.1605$ as the value of π ; but no trace has been found to prove the existence of demonstrative mathematics in ancient Egypt.

C. Mathematics in Mesopotamia

Sources abound for the study of Mesopotamian mathematics, and these sources may very well increase in the future. The Mesopotamians kept exact records of astronomical observations for long periods of time. Their more advanced mathematics was not limited to practical use, as was that of the Egyptians. They used a sexagesimal system of numeration with place value, and also used sexagesimal fractions; however, they lacked a cipher to denote zero until the 4th Century B.C., and they did not have a symbol corresponding to our decimal point, so that exact place value had to be determined from the context of each expression. They had a multiplication table and tables of inverses, squares, and cubes of numbers, and they used these tables to solve equations, even some simple equations of the third degree, as well as simultaneous equations of the second degree for two unknowns. They had accurate solutions for quadratic equations (expressed in words); they discarded negative roots, but they admitted both positive roots when two existed. They studied integral solutions of $a^2 + b^2 = c^2$ (the largest of their solutions were 12,709, 13,500, and 18,541) and approximate computation of quadratic roots, which suggests some relation to Greek mathematics. We have evidence that some of the geometric algebra in Euclid's *Elements* can be traced to Mesopotamian algebra. Some historians also affirm that the concept of demonstration in Greek mathematics originated with the Mesopotamians, but this theory lacks sufficient proof.

By the 7th Century, the Mayas in Central America also possessed a numeration system, with the base 20. As far as we know, the Mesopotamians and the Mayas were the earliest people to possess numeration systems with place value (- 57 Chinese Mathematics, 187 Greek Mathematics, 209 Indian Mathematics).

References

- [1] O. Neugebauer, Vorlesungen über Geschichte der antiken mathematischen Wissenschaften I, Springer, 1934.
- [2] O. Neugebauer, Mathematische Keilschrift-Texte, Springer, 1935.
- [3] A. B. Chace, The Rhind mathematical papyrus II, Mathematical Association of America, 1929.
- [4] B. L. van der Waerden, Science awakening, Noordhoff, 1954.
- [5] O. Neugebauer, The exact sciences in antiquity, Brown Univ. Press, second edition, 1957.

- [6] K. Vogel, Vorgriechische Mathematik I. Vorgeschichte und Ägypten II. Die Mathematik der Babylonier, Schroedel, 1958–1959.

25 (XX.35)

Approximation Methods in Physics

A. Introduction

It is often possible to solve a differential equation of mathematical physics analytically by expanding the solution in a series with respect to a small (or large) parameter involved in the equation. In general, such an approach is called a **perturbation method** (- Section B). For an equation which is difficult to treat analytically, it is possible to get a numerical solution either by replacing the derivatives by difference quotients ([†]difference method), or by expanding the solution in terms of suitable functions and determining the expansion coefficients numerically from the equation rewritten in a weak form ([†]variational method, [†]method of weighted residuals, [†]Galerkin's method, [†]finite-element method, etc.). For such numerical methods → 46 Calculus of Variations, 301 Numerical Solution of Algebraic Equations, 302 Numerical Solution of Linear Equations, 303 Numerical Solution of Ordinary Differential Equations, 304 Numerical Solution of Partial Differential Equations, 441 Variational Principles.

In Section C we describe a method for deriving an asymptotic expression for a function written as an integral in a complex plane.

B. Perturbation Method

(1) **Regular Perturbation.** (a) Initial-value problem. Consider an initial-value problem of an ordinary differential equation involving a small parameter ε :

$$\frac{du}{dt} = f(t, u; \varepsilon), \quad (1)$$

$$u = b(\varepsilon) \text{ at } t = 0. \quad (2)$$

We assume that f is sufficiently smooth as a function of t and is regular as a function of u and ε , and that b is a regular function of ε .

Substituting the power series in ε for f , u , and b ,

$$f(t, u; \varepsilon) = f^{(0)}(t, u) + \varepsilon f^{(1)}(t, u) + \dots$$

$$u(t; \varepsilon) = u^{(0)}(t) + \varepsilon u^{(1)}(t) + \dots$$

$$b(\varepsilon) = b^{(0)} + \varepsilon b^{(1)} + \dots$$

Approximation Methods in Physics

into equations (1) and (2) and equating terms of equal power, we get

$$\begin{aligned}\frac{du^{(0)}}{dt} &= f^{(0)}(t, u^{(0)}), \quad u^{(0)}(0) = b^{(0)}, \\ \frac{du^{(1)}}{dt} &= f_u^{(0)}(t, u^{(0)})u^{(1)} + f^{(1)}(t, u^{(0)}), \quad u^{(1)}(0) = b^{(1)}, \\ \frac{du^{(2)}}{dt} &= f_u^{(0)}(t, u^{(0)})u^{(2)} + \frac{1}{2}f_{uu}^{(0)}(t, u^{(0)})u^{(1)^2} \\ &\quad + f_u^{(1)}(t, u^{(0)})u^{(1)} + f^{(2)}(t, u^{(0)}), \\ &\quad u^{(2)}(0) = b^{(2)},\end{aligned}$$

the solutions of which determine the perturbation series for u . The equation for $u^{(0)}$ (the unperturbed equation) is nonlinear in general, but those for $u^{(1)}$, $u^{(2)}$, are inhomogeneous linear equations with identical principal parts.

The system of differential equations

$$\frac{du_i}{dt} = f_i(t, u_1, \dots, u_n; \varepsilon) \quad (i = 1, 2, \dots, n),$$

$$u_i = b_i(\varepsilon) \text{ at } t = 0,$$

or a differential equation of higher order,

$$\frac{d^n u}{dt^n} = f\left(t, u, \dots, \frac{d^{n-1} u}{dt^{n-1}}; \varepsilon\right),$$

$$u = b^{(0)}(\varepsilon), \dots, \frac{d^{n-1} u}{dt^{n-1}} = b^{(n-1)}(\varepsilon) \text{ at } t = 0,$$

can be treated in a similar way.

(b) Boundary-value problem. Given a linear differential operator $H = H_0 + \varepsilon V$ involving a small parameter ε linearly (H_0 is a linear differential operator and V is a function of (x_1, x_2, \dots, x_n)), consider the boundary-value problem

$$(H_0 + \varepsilon V)u = f(x_1, \dots, x_n) \quad (x_1, \dots, x_n) \in \Omega, \quad (3)$$

$$Bu = 0 \quad (x_1, \dots, x_n) \in \partial\Omega, \quad (4)$$

where B is a linear and homogeneous operator.

Substitution of the Taylor expansion

$$u = u_0 + \varepsilon u_1 + \dots$$

into equations (3) and (4) gives

$$H_0 u_0 = f, \quad Bu_0 = 0,$$

$$H_0 u_1 + V u_0 = 0, \quad Bu_1 = 0,$$

...

of which the solutions are

$$u_0 = Kf, \quad u_1 = -KVKf,$$

$$u_2 = KVKVKf, \dots$$

Here $K \equiv H_0^{-1}$ is an integral operator such that, for any function $v = v(x_1, x_2, \dots, x_n)$, we have

$$(Kv)(x_1, \dots, x_n) = \int_S \int_S G(x_1, \dots, x_n | \xi_1, \dots, \xi_n) \times v(\xi_1, \dots, \xi_n) d\xi_1 \dots d\xi_n,$$

where $G(x_1, \dots, x_n | \xi_1, \dots, \xi_n)$ is the Green's function for the unperturbed problem. The perturbation series

$$u = Kf - \varepsilon KVKf + \varepsilon^2 KVKVKf - \dots$$

converges if VK is bounded and $|\varepsilon| < 1/\|VK\|$.

(c) Eigenvalue problem. The solution of an eigenvalue problem

$$(H_0 + \varepsilon V)u = \lambda u \quad (5)$$

with the boundary condition $Bu = 0$ for the operators considered in (b) can be obtained in power series in ε :

$$\lambda = \lambda_0 + \varepsilon \lambda^{(1)} + \varepsilon^2 \lambda^{(2)} + \dots,$$

$$u = u_0 + \varepsilon u^{(1)} + \varepsilon^2 u^{(2)} + \dots,$$

where u_0 and λ_0 are the eigenfunction and nondegenerate eigenvalue, respectively, of the unperturbed operator H_0 , which is assumed to be self-adjoint. Substitution of the series into (5) gives

$$(H_0 - \lambda_0)u^{(1)} = -Vu_0 + \lambda^{(1)}u_0, \quad (6)$$

$$(H_0 - \lambda_0)u^{(2)} = -Vu^{(1)} + \lambda^{(2)}u_0 + \lambda^{(1)}u^{(1)}$$

$$(H_0 - \lambda_0)u^{(3)} = -Vu^{(2)} + \lambda^{(3)}u_0 + \lambda^{(2)}u^{(1)} + \lambda^{(1)}u^{(2)},$$

If we normalize u_0 and u by the conditions $(u_0, u_0) = 1$ and $(u, u_0) = 1$, we get from equation (6)

$$\lambda^{(1)} = (Vu_0, u_0) \equiv \alpha.$$

Let S be the inverse of $H_0 - \lambda_0 I$ in the subspace perpendicular to u_0 (we put $Su_0 = 0$), then each term of the perturbation series for u and λ is obtained in the following way from the system of the equations given above:

$$u^{(1)} = -SVu_0,$$

$$\lambda^{(2)} = -(VSvu_0, u_0),$$

$$u^{(2)} = S(V - \alpha)SVu_0,$$

$$\lambda^{(3)} = (VS(V - \alpha)SVu_0, u_0),$$

$$\begin{aligned}u^{(3)} &= -S(V - \alpha)S(V - \alpha)SVu_0 \\ &\quad + (VSvu_0, u_0)S^2Vu_0,\end{aligned}$$

(2) **Singular Perturbation.** Formal application of the procedure described in (1) often fails when the term including the highest-order derivative is multiplied by a small parameter (or a lower-order term is multiplied by a large parameter), i.e., when we deal with singular

perturbation problems. Typical methods for treating such cases are described below.

(a) **WKB method.** The method for getting an asymptotic solution of the second-order differential equation

$$\frac{d^2 u}{dx^2} + k^2 P(x)u = 0 \quad (7)$$

for large values of k was developed for problems arising in classical wave motion (Jeffreys, 1924 [1]) and for problems of quantum mechanics (Wentzel, Kramers, Brillouin, 1926), and is called the **Jeffreys method**, **WKB method**, or **WKB method**.

If the function P in (7) does not vary rapidly, the solution is expected to have the form

$$u \propto \exp\{ik\varphi(x)\},$$

with $\varphi(x)$ not very different from a linear function of x , since the exact solution for $P = P_0 = \text{const}$ is $\exp(\pm ik\sqrt{P_0}x)$.

Substitution of this expression for u in equation (7) gives

$$ik\varphi'' - k^2\varphi'^2 + k^2P = 0.$$

Neglecting the first term, we get

$$\varphi'^2 = P, \quad \text{i.e.,} \quad \varphi(x) = \int^x P^{1/2} dx.$$

This result suggests a transformation of the form

$$w = P^{1/4}u, \quad z = \int^x P^{1/2} dx,$$

which gives

$$\frac{d^2 w}{dz^2} + (k^2 - Q)w = 0, \quad (8)$$

where

$$Q = P^{-1/4} \frac{d^2}{dz^2} (P^{1/4}) = -P^{-3/4} \frac{d^2}{dx^2} (P^{-1/4}).$$

Equation (8) can be transformed further into a Volterra's integral equation of the second kind:

$$w(z; k) = Ae^{ikz} + Be^{-ikz} + \frac{1}{k} \int_{z_0}^z \sin\{k(z-\zeta)\} Q(\zeta) w(\zeta; k) d\zeta.$$

Therefore, if we expand w as

$$w(z; k) = w_0(z) + \frac{1}{k} w_1(z) + \dots, \quad (9)$$

we have

$$w_0(z) = Ae^{ikz} + Be^{-ikz},$$

$$w_n(z) = \int_{z_0}^z \sin\{k(z-\zeta)\} Q(\zeta) w_{n-1}(\zeta) d\zeta \quad (n=1, 2, \dots).$$

If P is positive for the interval $z_0 \leq z \leq z_1$ in question and has continuous derivatives up to second order, the foregoing series converges to the solution of (8).

On the other hand, when kz takes complex values, the series (9) does not converge in general. The series expansion appropriate for this case is

$$w(z; k) = \exp\{\psi(z; k)\}, \quad \psi = \pm ikz + \sum_{n=1}^{\infty} \frac{1}{(ik)^n} \psi_n(z),$$

and we get the asymptotic expansion for w :

$$w \sim \exp \left\{ \pm ikz \mp \frac{i}{2k} \int Q d\zeta + \frac{1}{4k^2} Q(z) \pm \frac{i}{8k^3} \left[Q'(z) - \int^z Q^2 d\zeta \right] + \frac{1}{16k^4} \{ Q''(z) - 2Q(z)^2 \} + \dots \right\}$$

Now, in many cases, P takes both positive and negative values within the interval in question, e.g., in the problem of transonic flow of a gas, in the problem of determining the reflectivity and transmissivity of a wave propagating in a nonuniform medium, or in the quantum-mechanical problem of a material wave passing through a potential wall. The point at which P vanishes is called a **turning point**. In this case there arises the problem of analytic continuation of the solution obtained in one side into the other through the turning point. For the special case when $P \propto x^n$ (the turning point is chosen as $x = 0$), it can be shown that the exact solution of equation (7) is explicitly given by

$$w \propto \sqrt{z} Z_v(kz),$$

where $v = 1/(n+2)$, and Z_v is a cylindrical function of order v . Based on this fact, the continuation formula for the case when P is approximated by x^n near $x = 0$ has been obtained.

(b) **Lighthill's method.** With nonlinear equations, formal application of a perturbation procedure sometimes leads to the difficulty that the solution of the unperturbed equation has a singularity not exhibited by the exact solution of the original equation and whose order increases in higher approximations, so that the perturbation series ceases to have any meaning near this point. Such a difficulty can often be overcome by so-called coordinate straining. This method has been applied to a number of fluid-dynamical problems, such as those arising in theories of aircraft wings, boundary layers, and shock waves.

Before describing the method, it is convenient to sketch **Poincaré's method** of getting the perturbation series for the period of a nonlinear oscillation. In order to solve the

nonlinear equation

$$\frac{du}{dt} = f(u; T_0, \varepsilon),$$

where T_0 is the period of oscillation when the small parameter ε is taken to be zero, Poincaré wrote series expansions for both u and the period T :

$$u = u_0(t) + \varepsilon u_1(t) + \dots, \quad T = T_0 + \varepsilon T_1 + \dots,$$

and determined T_1, T_2, \dots in such a manner as to avoid resonance at each step in the determination of $u_i(t)$, $u_i(t), \dots$

Based on this idea, Lighthill(1949) [2] developed coordinate straining in order to overcome the difficulty stated above. We illustrate the method here by means of the following simple example.

Consider a boundary-value problem of a nonlinear differential equation

$$(x + \varepsilon u) \frac{du}{dx} + u = 0 \quad (0 < x < 1): \quad (10)$$

$$u(1) = 1. \quad (11)$$

Application of the usual perturbation method would give the series solution

$$u = \frac{1}{x} + \frac{\varepsilon}{2} \left(\frac{1}{x} - \frac{1}{x^3} \right) - \frac{\varepsilon^2}{2} \left(\frac{1}{x^3} - \frac{1}{x^5} \right) + \dots$$

As is seen, however, the order of an apparent singularity occurring at $x = 0$ increases as we proceed to higher approximations so that it is not possible to see the true behavior of the solution near $x = 0$.

In Lighthill's method, we assume that series expansion is valid not only for u but also for the independent variable x , i.e.,

$$u = u_0(\xi) + \varepsilon u_1(\xi) + \dots, \quad x = \xi + \varepsilon x_1(\xi) + \dots$$

Substitution of these series into equations (10) and (11) gives

$$\xi u'_0 + u_0 = 0, \quad u_0(1) = 1,$$

$$\xi u'_1 + u_1 = -(u_0 x'_1 + u'_0 x_1 + u_0 u'_0),$$

$$u_1(1) = 0, \quad x_1(1) = 0,$$

From the first equation we have $u_0 = 1/\xi$. Before solving the second equation for u_1 , we determine $x_1(\xi)$ so as to avoid the increase of the order of singularity in $u_0(\xi)$ at $\xi = 0$. This requirement is satisfied by taking $x_1 = (1/2)(\xi - 1/\xi)$, and this gives $u_1 \equiv 0$; thus, at this stage of approximation, we have

$$u = u_0(\xi) + \varepsilon u_1(\xi) = \frac{1}{\xi},$$

$$x = \xi + \varepsilon x_1(\xi) = \xi + \frac{\varepsilon}{2} \left(\xi - \frac{1}{\xi} \right)$$

The guiding principle is the same for higher approximations. (The foregoing expression happens to be the exact solution, but this a fortuitous result of our choice of this particular example.)

The convergence of the series obtained by Lighthill's method has been proved by Wasow (1955) [3] for the more general problem

$$(x + \varepsilon u) \frac{du}{dx} + q(x)u = r(x),$$

$$u(a) = b,$$

where ε , a , and b are positive constants and $q(x)$ and $r(x)$ are functions regular in $|x| \leq a$.

Lighthill's method can also be applied to partial differential equations. Sometimes it is called the **PLK method (Poincaré-Lighthill-Kuo method)**, after its successful application to the problem of the boundary layer of a thin flat plate by Kuo.

(c) **Method of matched asymptotic expansions.** In some cases it is possible to get a uniformly valid series solution if we divide the domain into two or more subdomains, solve the equation by use of a suitable independent variable for each subdomain, and then determine the coefficients in each solution by the process of matching neighboring solutions on their common boundary. Such a perturbation technique was developed in treating the boundary-layer equation in fluid dynamics and has been systematized into the **method of matched asymptotic expansions**.

The idea of the method can be shown by the following simple example. Consider the boundary-value problem

$$\varepsilon \frac{d^2 u}{dx^2} + \frac{du}{dx} = a \quad (0 < x < 1),$$

$$u(0) = 0, \quad u(1) = 1, \quad (12)$$

where ε is a small positive parameter and a is a constant such that $0 < a < 1$. In order to solve this problem, we first take x itself to be an independent variable (the **outer variable**) in the domain $x \gg \varepsilon$ and expand the solution as

$$u = U_0(x) + \varepsilon U_1(x) +$$

Because the highest derivative is multiplied by a small parameter ε in equation (12), the perturbation equations to be solved are all of first order, of which the first one is

$$\frac{dU_0}{dx} = a, \quad U_0(1) = 1.$$

Solving this, we get an approximation for the **outer solution**:

$$U_0(x) = 1 - a(1 - x).$$

On the other hand, we introduce a new independent variable ξ (the **inner variable**) by

putting $x = \varepsilon \xi$ for $x \lesssim \varepsilon$. This transforms equation (12) into

$$\frac{d^2 v}{d\xi^2} + \frac{dv}{d\xi} = \varepsilon a, \quad v(\xi) \equiv u(x).$$

If we assume the expansion

$$v = V_0(\xi) + \varepsilon V_1(\xi) + \dots,$$

we have the equation for V_0 :

$$\frac{d^2 V_0}{d\xi^2} + \frac{dV_0}{d\xi} = 0, \quad V_0(0) = 0.$$

This gives an approximation for the **inner solution**:

$$V_0(\xi) = c(1 - e^{-\xi}).$$

Finally, in order to match the inner and the outer solutions, we equate their values at some point within the interval $\varepsilon \ll x \ll 1$ ($\xi \rightarrow \infty$ for the inner solution, $x \rightarrow 0$ for the outer solution):

$$V_0(\infty) = U_0(0) = 1 - a,$$

from which the coefficient in V_0 is determined to be $c = 1 - a$. Therefore we have an approximate solution valid in the whole interval $0 \leq x \leq 1$:

$$\begin{aligned} u &\cong U_0(x) + V_0(\xi) - (1 - a) \\ &= ax + (1 - a)(1 - e^{-x/\varepsilon}). \end{aligned}$$

C. Method of Steepest Descent

If an analytic function $f(z)$ of a complex variable z is expressed in terms of an analytic function $g(t)$ by an integral

$$f(z) = \int_C \exp\{zg(t)\} dt, \quad (13)$$

where C is a curve on the complex t -plane, then an asymptotic expression for $f(z)$ for large values of $|z|$ can be derived by the **method of steepest descent** (the **saddle-point method**). The idea of the method may be traced back to Riemann, and various asymptotic expressions for cylindrical functions were obtained by Debye (1909) [5].

The point at which the first derivative of $g(t)$ in the integral (13) vanishes is called a **saddle point**. The function $g(t)$ is expanded in a power series near this point as

$$g(t) = g(t_0) + \frac{1}{2}g''(t_0)(t - t_0)^2 + \dots$$

We have an inequality

$$zg''(t_0)(t - t_0)^2 \leq 0$$

along the line

$$L: \arg(t - t_0) = \frac{\pi}{2} - \frac{1}{2} \arg\{zg''(t_0)\}. \quad (14)$$

25 Ref.

Approximation Methods in Physics

Therefore the absolute value of the integrand $\exp\{zg(t)\}$ reaches a maximum at t_0 on the line L and decreases along it more rapidly than along any other direction. Hence if we deform the integration path in such a way that it passes through the saddle point t_0 in the direction of steepest descent, the value of the integrand is practically zero on the new path except very near to t_0 when $|z|$ is large, whereas the value of the integral remains the same. Therefore we can get an asymptotic expression of $f(z)$ for large $|z|$ by truncating the Taylor series for $g(t)$ up to the second term and taking the line L as the integration path:

$$\begin{aligned} f(z) &\approx \exp\{zg(t_0)\} \int_L \exp\left\{\frac{1}{2}zg''(t_0)(t - t_0)^2\right\} dt \\ &= \sqrt{\frac{2\pi}{-zg''(t_0)}} \exp\{zg(t_0)\}. \end{aligned}$$

As an example, **Stirling's formula**

$$n! \approx \sqrt{2\pi n} n^n e^{-n}$$

for a large positive integer is derived if we start with the formula

$$\begin{aligned} n! &= \int_0^\infty e^{-t} t^n dt = \int_0^\infty e^{-ns} (ns)^n d(ns) \\ &= n^{n+1} \int_0^\infty \exp\{n(\log s - s)\} ds \end{aligned}$$

and make use of the method of steepest descent.

References

- [1] H. Jeffreys, On certain approximate solutions of linear differential equations of the second order, London Math. Soc., (2) 23 (1924), 428-436.
- [2] M. J. Lighthill, A technique for rendering approximate solutions to physical problems uniformly valid, Phil. Mag., (7) 40 (1949), 117991201.
- [3] W. Wasow, On the convergence of an approximation method of M. J. Lighthill, J. Rational Mech. Anal., 4 (1955), 751-767.
- [4] H. S. Tsien, The Poincaré-Lighthill-Kuo method, Advances in Applied Mechanics, Academic Press, 1956, vol. 4, 281-349.
- [5] P. Debye, Näherungsformeln für die Zylinderfunktionen für grosse Werte des Arguments und unbeschränkt veränderliche Werte des Index, Math. Ann., 67 (1909), 535-558.
- [6] P. M. Morse and H. Feshbach, Methods of theoretical physics, McGraw-Hill, 1953.
- [7] H. Jeffreys and B. S. Jeffreys, Methods of mathematical physics, Cambridge Univ. Press, third edition, 1962.
- [8] J. D. Cole, Perturbation methods in applied mathematics, Blaisdell, 1968.

[9] A. H. Nayfeh, *Perturbation methods*, Wiley, 1973.

[10] A. Erdélyi, *Asymptotic expansions*, Dover, 1956.

26 (Xx1.4) Arab Mathematics

The role of the Arabs in cultural history has been partly that of cultural transmitter. Between the 7th and 13th centuries, they established a religious empire that extended from India to Spain; later it was divided into the eastern and western empires. The caliphs of these empires encouraged research in the sciences, so the capitals Baghdad and Cordova became centers of culture where scholars from different countries gathered.

Arabic scholarship is sometimes called the stepfather of European culture. During the 13th Century, Alphonso X (1252–1284) invited Islamic and Hebrew scholars to the Spanish court to translate their writings on algebra, medicine, and astronomy into Spanish. This accomplishment earned him the title of Alphonso the Wise.

The first contact between Greek and Indian mathematics took place in Baghdad under Caliph Al-Mansūr (754–775); Euclid's *Elements* was introduced by way of the Byzantine Empire, while Brahmagupta's *Brahmasphutasiddhānta* came directly from India. Many mathematical texts found in the Eastern Roman Empire and Syria, including some Greek works, were translated into Arabic. Though it is difficult to discern essential scientific advances in Arabian works, the diffusion of these translations was instrumental in the development of European mathematics.

The Arabs did not use written numerals until Mohammed's time (570–632). Signs representing numbers had been introduced into Arabia when its influence encompassed Egypt and Greece. Indian numerals were imported with Brahmagupta's book and became our present **Arabic numerals** after a series of modifications.

Among all the branches of Arab mathematics, algebra was the most advanced. It started with Alkwarizmi's (820) *Al Gebr W'al Muqabala*, the origin of the word "algebra." It was the first mathematical book written in Arabic. Its content was essentially a variety of methods of solving algebraic equations. *Al gebr* means "transposition of negative terms on one side of the equation to the other side and changing their signs," and *al muqabala* means "simplification of the equation by

gathering similar terms." For example, quadratic equations can be brought by these methods into one of the following three types: $x^2 = px + q$; $x^2 + q = px$; $x^2 + px = q$, where p, q are positive numbers. The Arabs expressed the rule of solving these equations verbally. They apparently knew that quadratic equations have two roots, but they adopted only positive roots; when the equation had two positive roots they adopted the smaller root. The proof was given geometrically. It is possible that they learned geometric proofs from the Greeks.

For the Arabs, geometry was secondary to algebra. They did not appreciate proof as seen in Euclid's *Elements*. The book on conic sections by Apollonius was also translated into Arabic, but no essential progress was made in this area. The only remarkable contribution was that of Omar Khayyám, author of *The Rubáiyát*, who applied conic sections to the solution of the cubic equation $x^3 + bx = a$.

In trigonometry, Al Battani (c. 858–929) left a notable contribution. He studied *The Almagest*, the Arabic translation of Ptolemy's astronomical work. He added nothing outstanding to plane trigonometry, but obtained such formulas as $\cos a = \cos b \cos c + \sin b \sin c \cos \alpha$ for spherical triangles, which were not mentioned in *The Almagest*.

References

- [1] M. Cantor, *Vorlesungen über Geschichte der Mathematik I*, Teubner, third edition, 1907.
- [2] G. Sarton, *Introduction to the history of science I. From Homer to Omar Khayyam*, Carnegie Institution of Washington, 1927.

27 (V.18) Arithmetic of Associative Algebras

A. General Theory

Let \mathfrak{g} be a \dagger Dedekind domain (i.e., an integral domain in which every ideal is uniquely decomposed into a product of prime ideals), let F be the field of quotients of \mathfrak{g} , and let A be a \dagger separable algebra of finite degree over F . A \mathfrak{g} -lattice \mathfrak{a} of A is a \mathfrak{g} -submodule of A that is finitely generated over \mathfrak{g} and satisfies $A = F\mathfrak{a}$. If a subring \mathfrak{o} of A is a \mathfrak{g} -lattice containing \mathfrak{g} , then \mathfrak{o} is called an **order**. A **maximal order** is an order that is not contained in any other order. A maximal order always exists although

it may not be unique; in particular, if A is commutative, it has only one maximal order. If we put $\mathfrak{o}_l = \{x \in A \mid xa \subset \mathfrak{a}\}$ ($\mathfrak{o}_r = \{x \in A \mid ax \subset \mathfrak{a}\}$) for a g -lattice \mathfrak{a} of A , then \mathfrak{o}_l (\mathfrak{o}_r) is an order of A and is called the **left (right) order** of \mathfrak{a} . If the left order of \mathfrak{a} is maximal, then so is the right order; the converse is also true. If \mathfrak{o}_l and \mathfrak{o}_r are maximal, we call \mathfrak{a} a **normal g -lattice**. To describe the same situation we say that \mathfrak{a} is a **left \mathfrak{o}_l -ideal** or a **right \mathfrak{o}_r -ideal**. If $\mathfrak{o}_l = \mathfrak{o}_r = \mathfrak{o}$, then we say that \mathfrak{a} is a **two-sided \mathfrak{o} -ideal**. A lattice \mathfrak{a} with $\mathfrak{a} \subset \mathfrak{o}_l$ (or equivalently with $\mathfrak{a} \subset \mathfrak{o}_r$) is called an **integral g -lattice** or **integral left (right) ideal**. The product ab of two normal g -lattices $\mathfrak{a}, \mathfrak{b}$ is called a **proper product** if the right order of \mathfrak{a} coincides with the left order of \mathfrak{b} . The proper product defines the structure of a g -groupoid on the set of all normal g -lattices of A . In particular, the inverse $^{-1}$ of a normal g -lattice \mathfrak{a} satisfying $\mathfrak{a}\mathfrak{a}^{-1} = \mathfrak{o}_l$, $\mathfrak{a}^{-1}\mathfrak{a} = \mathfrak{o}_r$ is given by $\mathfrak{a}^{-1} = \{x \in A \mid xa \subset \mathfrak{o}_r\} = \{x \in A \mid ax \subset \mathfrak{o}_l\}$. For a fixed maximal order \mathfrak{o} , a maximal **integral two-sided \mathfrak{o} -ideal** \mathfrak{p} different from \mathfrak{o} is called a **prime ideal** of \mathfrak{o} . If \mathfrak{p} is prime, then $\mathfrak{o}/\mathfrak{p}$ is the matrix algebra of degree κ over a division algebra, and κ is called the **capacity** of the prime ideal \mathfrak{p} . The set of all two-sided \mathfrak{o} -ideals forms a multiplicative group, of which prime ideals are independent generators.

B. Maximal Orders of a Simple Ring

In the rest of this article, A is a \dagger simple ring, F is the \dagger center of A , and \mathfrak{o} is a maximal order of A . The prime ideals \mathfrak{q} of \mathfrak{o} and the prime ideals \mathfrak{p} of g are in one-to-one correspondence by the relation $\mathfrak{q} \cap g = \mathfrak{p}$; $\mathfrak{o}/\mathfrak{q}$ is a simple algebra over g/\mathfrak{p} , and $\mathfrak{q}^e = \mathfrak{p}\mathfrak{o}$ for some natural number e . The **different** \mathfrak{d} of \mathfrak{o} is defined by $\mathfrak{d}^{-1} = \{x \in A \mid \text{Tr}(x\mathfrak{o}) \subset g\}$. (Tr is the reduced trace from A to F ; \mathfrak{d} is an integral two-sided \mathfrak{o} -ideal, and is divisible by \mathfrak{q}^{e-1} if $\mathfrak{q}^e = \mathfrak{p}\mathfrak{o}$.) For \mathfrak{q} to divide \mathfrak{d} , it is necessary and sufficient that either $e > 1$ or $\mathfrak{o}/\mathfrak{q}$ not be separable over g/\mathfrak{p} . In particular, if A is a \dagger total matrix algebra over F , then $\mathfrak{d} = \mathfrak{o}$, $\mathfrak{q} = \mathfrak{p}\mathfrak{o}$. The ideal of g generated by the \dagger reduced norms (to F) of the elements in a normal g -lattice \mathfrak{a} of A is denoted by $N_{A/F}(\mathfrak{a})$. If ab is a proper product, then we have $N_{A/F}(ab) = N_{A/F}(\mathfrak{a})N_{A/F}(\mathfrak{b})$, where $N_{A/F}(\mathfrak{d})$ does not depend on the choice of \mathfrak{o} ; this is called the **discriminant** of A . If $[A : F] = n^2$ and $\mathfrak{q}^e = \mathfrak{p}\mathfrak{o}$, then $N_{A/F}(\mathfrak{q}) = \mathfrak{p}^f$, $ef = n$.

C. Simple Rings over a Local Field

Let F be a field that is complete with respect to a \dagger discrete valuation whose field of residue classes is finite. Let g be the \dagger valuation ring of

F , \mathfrak{p} be the maximal ideal of g , and A be a simple algebra with F as its center. If A is a \dagger division algebra and $[A : F] = n^2$, then A has only one maximal order \mathfrak{o} , and \mathfrak{o} has only one prime ideal \mathfrak{q} such that $\mathfrak{q}^n = \mathfrak{p}\mathfrak{o}$; $\mathfrak{o}/\mathfrak{q}$ is an extension of degree n of g/\mathfrak{p} . If A is not necessarily a division algebra, the relation $\xi\mathfrak{o}\xi^{-1} = \mathfrak{o}'$ with an element ξ of A holds between two maximal orders \mathfrak{o} and \mathfrak{o}' of A . Furthermore, for any left (right) \mathfrak{o} -ideal \mathfrak{a} , there exists an element α such that $\mathfrak{a} = \mathfrak{o}\alpha$ ($\mathfrak{a} = \alpha\mathfrak{o}$). By using the notation of \dagger cyclic algebra, we can express A in the form (K_n, σ, π) . Here K_n is the unramified extension of degree n over F , σ is the \dagger Frobenius substitution of K_n/F , π is a prime element of F , and $0 \leq \gamma < n$. The element of \mathbf{Q}/\mathbf{Z} determined by $\gamma/n \pmod{\mathbf{Z}}$ is denoted by $\langle A \rangle$ and is called the \dagger Hasse invariant of the algebra class containing A . The mapping $A \mapsto \langle A \rangle$ gives an isomorphism of the \dagger Brauer group of F onto the additive group \mathbf{Q}/\mathbf{Z} . If M is an extension of F of finite degree, then $\langle A^M \rangle = [M : F] \langle A \rangle$ holds for the algebra A^M obtained from A by scalar extension (- 29 Associative Algebras).

D. Simple Rings over an Algebraic Number Field

Let F be an algebraic number field of finite degree, and let A be a simple algebra with center F . Then A is a cyclic algebra and is isomorphic to a total matrix algebra over a division algebra D . The order n of the algebra class of A over F is determined by $n^2 = [D : F]$ (H. Hasse, R. Brauer, and E. Noether).

Denote by $F_{\mathfrak{p}}$ the completion of F with respect to a \dagger prime divisor \mathfrak{p} of F , and let $A_{\mathfrak{p}}$ be the algebra obtained from A by the scalar extension $F_{\mathfrak{p}}$ over F . For a finite prime divisor \mathfrak{p} , the meaning of $\{A_{\mathfrak{p}}\}$ is as before; for an infinite prime divisor \mathfrak{p} , put $\{A_{\mathfrak{p}}\} = 0$ or $1/2 \pmod{\mathbf{Z}}$ according as $A_{\mathfrak{p}}$ is a total matrix algebra over $F_{\mathfrak{p}}$ or not. Furthermore, define the subgroup $J_{\mathfrak{p}}$ of \mathbf{Q}/\mathbf{Z} by

$$\begin{aligned} J_{\mathfrak{p}} &= \mathbf{Q}/\mathbf{Z}, & \mathfrak{p} \text{ a finite prime divisor,} \\ &= \{0, 1/2 \pmod{\mathbf{Z}}\} & \mathfrak{p} \text{ a real infinite prime divisor,} \\ &= \{0\}, & \mathfrak{p} \text{ a complex infinite prime divisor.} \end{aligned}$$

Now let J be the subgroup of the direct product $\prod_{\mathfrak{p}} J_{\mathfrak{p}}$ consisting of all elements of the form $(\alpha_{\mathfrak{p}})$ ($\alpha_{\mathfrak{p}} \in J_{\mathfrak{p}}$) such that $\alpha_{\mathfrak{p}} = 0$ except for a finite number of prime divisors and $\sum_{\mathfrak{p}} \alpha_{\mathfrak{p}} = 0$. Then $A + (\{A_{\mathfrak{p}}\})$ gives rise to an isomorphism of the Brauer group over F onto J (Hasse). Each $\langle A_{\mathfrak{p}} \rangle$ is called the \dagger \mathfrak{p} -invariant of A . In particular, A is a total matrix algebra over F if

and only if A , is a total matrix algebra over F_p for all p (- 29 Associative Algebras). These theorems are closely related to †class field theory.

If \mathfrak{o} is a maximal order of A and a, b are left \mathfrak{o} -ideals, then $a\xi = b$ with an element ξ of A defines an equivalence relation between a and b . The number of equivalence classes of left ideals with respect to this equivalence is called the **class number** of A ; it is independent of the choice of \mathfrak{o} and is equal to the class number defined by using right ideals. The product of all real infinite prime divisors p with $\{A_p\} = 1/2$ is denoted by P_∞ . If P_∞ is the product of all infinite prime divisors and $[A : F] = 4$, then A is called a **totally definite quaternion algebra**.

If \mathfrak{o} is a maximal order of A and A is not a totally definite quaternion algebra, then $\mathfrak{a} \rightarrow N_{\infty}(\mathfrak{a})$ gives a one-to-one correspondence between the classes of left \mathfrak{o} -ideals and the congruence classes of ideals of F modulo P_∞ (- 14 Algebraic Number Fields H) (Eichler's theorem).

In particular, if A is a total matrix algebra over F , then the class number of A is equal to the class number of F . The class number of a totally definite quaternion algebra was determined by M. Eichler by using the zeta function (- Section F) of A [4].

Let \mathfrak{o} be a maximal order of A , a be an integral two-sided \mathfrak{o} -ideal, b be an integer of F , and ξ be an element of \mathfrak{o} . Furthermore, assume $b \equiv 1 \pmod{P_\infty}$, $N_{A/F}(\xi) \equiv b \pmod{a \cap F}$ (†multiplicative congruence). Then there exists an element β of \mathfrak{o} such that $N_{A/F}(\beta) = b$, $\beta \equiv \xi \pmod{a}$ ($N_{A/F}$ is the reduced norm), provided that A is not a totally definite quaternion algebra [5]. This theorem, which is called **Eichler's approximation theorem**, is widely applicable; e.g., it yields the previous theorem on the class number and can be generalized to the case of semisimple †algebraic groups (- 13 Algebraic Groups).

E. Algebras over a Function Field

The Hasse-Brauer-Noether and Hasse theorems also hold for †normal simple algebras over a †field of algebraic functions of one variable over a finite field. On the other hand, a normal simple algebra over a field K of algebraic functions of one variable over an algebraically closed field is a total matrix algebra over K (Tsen's theorem).

F. Other Notions

Adeles and ideles for a simple algebra A over an algebraic number field of finite degree can

be introduced as in the case of number fields (\rightarrow 6 Adeles and Ideles). If $N(\mathfrak{a})$ stands for the number of elements in $\mathfrak{o}/\mathfrak{a}$, where \mathfrak{o} is a maximal order of A and \mathfrak{a} is an integral left \mathfrak{o} -ideal, then the **zeta function of the simple algebra A** , called the **Hey zeta function**, is defined by $\zeta_A(s) = \sum N(\mathfrak{a})^{-s}$ (the sum over all integral left \mathfrak{o} -ideals). This function has properties similar to those of Dedekind zeta functions (- 450 Zeta Functions L). Let p be an infinite prime divisor of the center F of A , and let G_p be the group of elements in A , with the reduced norm 1. Put $G = \prod_p G_p$ (the product over all infinite prime divisors of F). Then the group Γ of units with the reduced norm 1 in \mathfrak{o} is naturally regarded as a subgroup of G . To be more precise, Γ is a discrete subgroup of G , the volume of G/Γ is finite with respect to an invariant measure, and G/Γ is compact if and only if A is a division algebra (\rightarrow 122 Discontinuous Groups). This result can be viewed as a special case of more general facts about semisimple algebraic groups (- 13 Algebraic Groups). If K is a †maximal compact subgroup of G , then Γ gives rise to a †discontinuous group operating on the homogeneous space G/K , and we obtain †automorphic forms with respect to Γ . The case where A is a †quaternion algebra has been studied extensively (- 32 Automorphic Functions).

References

- [1] C. Chevalley, L'arithmétique dans les algèbres de matrices, Actualités Sci. Ind., Hermann, 1936.
- [2] M. Deuring, Algebren, Erg. Math., Springer, second edition, 1968.
- [3] M. Eichler, Über die Idealklassenzahl hyperkomplexer Systeme, Math. Z., 43 (1938), 481–494.
- [4] M. Eichler, Über die Idealklassenzahl total definiter Quaternionen Algebren, Math. Z., 43 (1938), 102–109.
- [5] M. Eichler, Allgemeine Kongruenzklasseneinteilungen der Ideale einfacher Algebren über algebraischen Zahlkörpern und ihre L -Reihen, J. Reine Angew. Math., 179 (1938), 227–251.
- [6] N. Jacobson, The theory of rings, Amer. Math. Soc. Math. Surveys, 1943.
- [7] C. L. Siegel, Discontinuous groups, Ann. Math., (2) 44 (1943), 674–689. (Gesammelte Abhandlungen, Springer, 1966, vol. 2, 390–405.)
- [8] A. Weil, Adeles and algebraic groups, Lecture notes, Institute for Advanced Study, Princeton, 1961.

28 (XXI.1 3) Artin, Emil

Emil Artin (March 3, 1898–December 20, 1962) was born in Vienna. After he studied at the Universities of Vienna, Leipzig, and Göttingen, he taught at the University of Hamburg from 1923 to 1937. In 1937 he left Germany under the Nazi regime for America, where he taught at Notre Dame, Indiana, and Princeton universities. He returned to Germany in 1948 and taught again at the University of Hamburg until his fatal heart attack.

In his thesis (1923), he proved affirmatively the \dagger Riemann hypothesis in function fields over finite constant fields in the hyperelliptic case, and conjectured that this should be valid in general. This was eventually verified by A. Weil in 1941. Also in 1923, he introduced the \dagger L-function for Galois extension and was led to the \dagger general law of reciprocity, which he stated as a conjecture and proved four years later, thus bringing \dagger class field theory to completion. Around the same period, he established in collaboration with O. Schreier the theory of \dagger formally real fields and solved the 17th \dagger problem of Hilbert. He also initiated the theory of \dagger braids, which he later developed with F. Bohnenblust in the 1940s. In view of these and other ingenious works as well as his inspiring teaching in all areas of mathematics, Artin is considered to have been one of the most influential mathematicians of this Century.

References

- [1] S. Lang and J. T. Tate (eds.), The collected papers of Emil Artin, Addison-Wesley, 1965.
- [2] H. Cartan, Emil Artin, Hamb. Abh., 29 (1965).

29 (III.1 0) Associative Algebras

A. Fundamental Concepts

Let K be a commutative ring with unity element 1 (- 368 Rings A), and let A be a ring which is a \dagger unitary K -module (- 277 Modules). Such a ring A is called an **associative algebra over K** (or simply **algebra over K**) if it satisfies the condition $\lambda(ab) = (\lambda a)b = a(\lambda b)$ ($\lambda \in K$; $a, b \in A$). An (associative) algebra A over

K is often written A/K , and K is called the **coefficient ring** (or **ground ring**) of the algebra $A = A/K$. In particular, if K is a \dagger field, then it is called the **coefficient field** (or **ground field**) of A . Algebras over fields have been studied in detail. Notions such as **zero algebra**, **unitary algebra**, **commutative algebra**, **(semi) simple algebra**, and **division algebra** are replicas of the respective ones for rings (- 368 Rings). Considering both structures as rings and as K -modules, homomorphisms and isomorphisms are defined in a natural manner, and are called **algebra homomorphism** and **algebra isomorphism**, respectively. In this connection, **subalgebra**, **quotient algebra** (or **residue class algebra**), and **direct product** of algebras are also defined as in the case of rings.

An \dagger **ideal** of an algebra A is defined as an ideal of the ring A , which is at the same time a submodule of the module A over the coefficient ring. The \dagger radical of an algebra A considered as a ring is then an ideal of A in this sense. In fact, the existence of a unity in an algebra A implies that any ideal of A considered as a ring is necessarily an ideal of the algebra A .

In the rest of this article, we assume that all rings have a unity element and that all homomorphisms are unitary. Hence, when we consider subalgebras of an algebra A , we require that they share the unity element with A . If e is the unity element of an algebra A over K , then the mapping $\lambda \rightarrow \lambda e = \lambda'$ ($\lambda \in K$) is a homomorphism $K \rightarrow A$ whose image Ke is contained in the \dagger center of A , and the scalar multiplication λa is equal to the ring multiplication $\lambda' a$ ($\lambda \in K$, $a \in A$). Conversely, given a homomorphism of K into a ring A whose image is contained in the center of A , we can regard A as an algebra over K in an obvious way. Hence we are given an algebra A over K if and only if there exists a pair (A, ρ) of a ring A and a homomorphism $\rho: K \rightarrow A$ whose image is contained in the center of A . There exists a uniquely determined (unitary) homomorphism of the ring \mathbb{Z} of rational integers into any ring; hence any ring can be regarded as an algebra over \mathbb{Z} . If the coefficient ring K of a nonzero algebra A is a field, then K can be regarded as a subfield contained in the center of A , and the unity element 1 of K coincides with the unity element of A .

Let A and B be algebras over K . Then the \dagger tensor product $A \otimes_K B$ of K -modules is an algebra over K under the multiplication $(a \otimes b)(a' \otimes b') = aa' \otimes bb'$ ($a, a' \in A$, $b, b' \in B$). This algebra is called the **tensor product** of algebras A and B . Moreover, the mapping $a \rightarrow a \otimes 1$ (resp. $b \rightarrow 1 \otimes b$) ($a \in A$, $b \in B$) gives an algebra homomorphism $A \rightarrow A \otimes_K B$ (resp. $B \rightarrow$

$A \otimes_K B$, which is called the **canonical homomorphism**. In particular, if A is a commutative algebra, then $A \otimes_K B$ can be regarded as an algebra over A by this canonical homomorphism; and in this case $A \otimes_K B$ is called the algebra obtained by **extension of the coefficient ring** of B to A (or a **scalar extension** of B by A), and is often denoted by B^A . The algebra $K \otimes_K B$ is canonically isomorphic to B . Furthermore, $(A \otimes_K B) \otimes_K C$ and $A \otimes_K (B \otimes_K C)$ are canonically isomorphic and are written $A \otimes_K B \otimes_K C$. Let A and B be commutative algebras over K . Then, for any commutative algebra C and homomorphisms $\alpha: A \rightarrow C$, $\beta: B \rightarrow C$, there exists one and only one homomorphism $\gamma: A \otimes_K B \rightarrow C$ such that $\alpha(a) = \gamma(a \otimes 1)$, $\beta(b) = \gamma(1 \otimes b)$ ($a \in A$, $b \in B$). This property characterizes the tensor product $A \otimes_K B$ of commutative algebras A and B . In this sense, $A \otimes_K B$ is sometimes called the **coproduct** of A and B (- 52 Categories and Functors E).

B. Examples of Associative Algebras

As we mentioned, any ring can be regarded as an algebra over the ring \mathbb{Z} of rational integers. But it is often useful to deal with algebras over "large" or more "efficient" coefficient rings. For instance, we have many rings which are algebras over a commutative ring K , such as the ring of polynomials, the ring of formal power series in n variables with coefficients in K , the endomorphism ring of a K -module, and the full matrix ring of degree n over K (- 368 Rings C). There are other important classes of algebras, such as (semi) group algebras, Hecke algebras, and crossed-product algebras, which will be explained later. These algebras are defined by a canonical basis connected directly with a (semi) group structure. On the other hand, the tensor algebra and the exterior algebra of linear spaces and the Clifford algebra associated with a given quadratic form are also important (- 61 Clifford Algebras, 256 Linear Spaces).

The most frequently used example of a division algebra is the **quaternion field \mathbf{H}** (often called **Hamilton's quaternion algebra**, W. R. Hamilton, 1858). This is a 4-dimensional linear space over the real number field \mathbf{R} with basis $\{1, i, j, k\}$, with the following laws of multiplication: 1 is a unity element, $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, and $ki = -ik = j$. An element of \mathbf{H} is called a **quaternion**. The only finite-dimensional division algebras over the real number field \mathbf{R} are the real number field \mathbf{R} , the complex number field \mathbf{C} , and the quaternion field \mathbf{H} .

C. Group Algebras and Hecke Algebras

Let K be a commutative ring and $K^{(G)}$ be the direct sum $\sum_{s \in G} K_s$ of modules K_s , where each K_s is isomorphic to K with a group G as the index set (- 277 Modules F). The elements of $K^{(G)}$ are the families $(\lambda_s)_{s \in G}$ of elements of K whose components are all zero except for a finite number of them. Let $\{u_s\}_{s \in G}$ be the canonical basis of $K^{(G)}$, namely, u_s is an element of $K^{(G)}$ of which the s th component is 1 and the others are 0. The module $K^{(G)}$ has the structure of an algebra over K , where the law of multiplication is determined by $u_s u_t = u_{st}$ ($s, t \in G$). This algebra is called the **group algebra** of G over K . The product $\lambda * \mu$ of elements $\lambda = (\lambda_s)$ and $\mu = (\mu_s)$ of $K^{(G)}$ is then given by

$$(\lambda * \mu)_s = \sum_{s=rt} \lambda_r \mu_t, \quad s \in G. \quad (1)$$

Each basis element u_s is often identified with the group element s , and in this manner, the group G is regarded as a basis of $K^{(G)}$, which is usually written KG or $K[G]$.

In this definition, the group G can be replaced by a semigroup G , and then the algebra $K^{(G)}$ is called a **semigroup algebra**. As an example, let \mathbf{N} be the additive semigroup of all nonnegative rational integers and \mathbf{N}^n be the direct product of n copies of \mathbf{N} . If we denote the elements (i_1, \dots, i_n) of \mathbf{N}^n by $X_1^{i_1} \dots X_n^{i_n}$ and use multiplication instead of addition, then the semigroup algebra of \mathbf{N}^n over K is exactly the ring of polynomials $K[X_1, \dots, X_n]$. On the other hand, if G is a semigroup, then even when it is infinite, it may occur that for any $s \in G$, there exists only a finite number of pairs (r, l) of elements of G such that $s = rl$. In this case, formula (1) also defines the law of multiplication on the Cartesian product K^G . This algebra is called a **large semigroup algebra** and contains $K^{(G)}$ as a subalgebra. In particular, the large semigroup algebra of \mathbf{N}^n is exactly the ring of formal power series $K[[X_1, \dots, X_n]]$.

Let H be a subgroup of a group G , and assume that the index of $H \cap sHs^{-1}$ in H is finite for any $s \in G$. This assumption is equivalent to the condition that any double coset of G by H is a union of a finite number of left as well as right cosets. Let $H \backslash G$, G/H , and $H \backslash G/H$ be the set of all right cosets, left cosets, and double cosets, respectively. Then each element of the direct sum $K^{(H \backslash G)}$, $K^{(G/H)}$, or $K^{(H \backslash G/H)}$ can be regarded as a function defined on G taking a constant value on each right, left, or double coset, respectively. Conversely, any function defined on G can be regarded as an element of $K^{(H \backslash G)}$, $K^{(G/H)}$, or

$K^{(H \setminus G/H)}$ if it takes a constant value on each left, right, or double coset, respectively, and if it vanishes everywhere except on a finite number of cosets of respective type. Under such identification, let $\lambda_s, \lambda_l, \lambda_r,$ and λ_t denote the values of a function $\lambda: G \rightarrow K$ on $s \in G, l \in H \setminus G, r \in G/H,$ and $t \in H \setminus G/H,$ respectively. For any $\lambda, \mu \in K^{(H \setminus G/H)}$ we define a function $\lambda * \mu$ by

$$(\lambda * \mu)_s = \sum_{s \in rl} \lambda_r \mu_l, \quad s \in G, \quad (2)$$

where the right-hand side is the sum taken over all pairs (r, l) such that $s \in rl, r \in G/H, l \in H \setminus G$. It can be shown that this sum is a finite sum and that $\lambda * \mu \in K^{(H \setminus G/H)}$. Hence the module $K^{(H \setminus G/H)}$ has the structure of an algebra over K , which is called the **Hecke algebra** of (G, H) over K and often denoted by $\mathcal{H}_K(G, H)$. If $H = \{e\}$, the $\mathcal{H}_K(G, H)$ is exactly the group algebra of G . In general, $\mathcal{H}_K(G, H)$ can be regarded as an algebra obtained by extending the coefficient ring of $\mathcal{H}_Z(G, H)$ to K . Furthermore, $K^{(H \setminus G)} (K^{(G/H)})$ can be regarded as a right (left) module over the group algebra $K^{(G)}$, and the endomorphism ring $\mathcal{E}_{K^{(G)}}(K^{(H \setminus G)})$ (resp. $\mathcal{E}_{K^{(G)}}(K^{(G/H)})$) is canonically isomorphic (anti-isomorphic) to the Hecke algebra $K^{(H \setminus G/H)}$.

D. General Crossed Product

Let G be a group that operates on a commutative ring L , and denote the operation by $(s, \lambda) \rightarrow s(\lambda)$ ($s \in G, \lambda \in L$); thus for any $s \in G$ the mapping $\lambda \rightarrow s(\lambda)$ ($\lambda \in L$) is an automorphism of L satisfying $s(t(\lambda)) = st(\lambda)$ ($s, t \in G$). For any $\lambda, \mu \in L^{(G)}$, we define the product $\lambda * \mu \in L^{(G)}$ by

$$(\lambda * \mu)_s = \sum_{s \in rl} \lambda_r \mu_l f(r, l), \quad s \in G, \quad (3)$$

where $\{f(r, l)\}_{r, l \in G}$ is a given family of elements of L . If this family satisfies the equations

$$f(s, r)f(sr, l) = s(f(r, l))f(s, rl), \quad s, r, l \in G,$$

then $L^{(G)}$ forms a ring. In terms of the canonical basis $\{u_s\}_{s \in G}$, this ring structure is defined by the formulas $u_r u_l = f(r, l) u_{rl}, u_s \lambda = s(\lambda) u_s$ ($\lambda \in L$). If K is the subring of L consisting of all $\lambda \in L$ such that $s(\lambda) = \lambda$ ($s \in G$), then the ring $L^{(G)}$ is an algebra over K , called the **crossed product** of L and G with respect to the given operation and the given **factor set** f of G . In a narrower sense of the term, we consider only the case when L is a field, G is a finite group, $f(r, l) \neq 0$, and G operates on L faithfully. In this case, G can be identified with the t Galois group of a finite t Galois extension L/K , and so the crossed product is written $(L/K, f)$. This is a t central simple algebra over K (- Sections E, F).

If the operation of G in the above-mentioned general crossed product is trivial,

namely, $L = K$, the crossed product of K and G is called an **algebra extension** of G over K with respect to f . Usually, we assume that K is a field and $f(r, l) \neq 0$. If $f(r, l) = 1$, the algebra extension is the group algebra. If G is a finite group whose order is prime to the t characteristic of K , then any algebra extension of G over K (in particular, the group algebra) is always t semisimple and t separable. If G is a finite Abelian group and f is a factor set of G , then a bihomomorphism $\varphi: G \times G \rightarrow K^*$ (a mapping which is a homomorphism in each variable) is defined by $\varphi(s, t) = f(s, t) \cdot f(t, s)^{-1}$. An algebra extension of G over K with respect to f is a central simple algebra if and only if φ is nondegenerate. In particular, if G is a direct product of n copies of a group of order 2, then by choosing n elements s_1, s_2, \dots, s_n of G suitably, any element of G can be uniquely expressed as $s_a s_b \dots s_z$, where $\{a, b, \dots, z\}$ is a subset of $\{1, \dots, n\}$ in the natural order. Hence, if we take $u_{a, \dots, z} = u_a u_b \dots u_z$ as a basis of an algebra extension, then any factor set f is determined by $f(s_i, s_j) = \lambda_{ij}$, where $\lambda_{ij} = 1$ ($i < j$), $\lambda_{ij} = \pm 1$ ($i > j$), and λ_{ii} are arbitrary. When $\lambda_{ij} = -1$ ($i > j$), the corresponding algebra extension is a t Clifford algebra. Furthermore, if $\lambda_{ii} = 0$, then it is a t Grassmann algebra. If $\lambda_{ii} \neq 0$ and n is even, then it is a central simple algebra whenever the characteristic of K is not 2 (- 61 Clifford Algebras).

Now let K be the real number field \mathbf{R} and A , be the Clifford algebra with $\lambda_{ii} = 1$. Then A , is the quaternion field. Elements of A , are called **sedentions**, and are important in t spinor theory and t Dirac's equation. In general, let K be an arbitrary field whose characteristic is not 2 and put $n = 2, \lambda_{12} = 1, \lambda_{21} = -1, \lambda_{11} = \lambda \neq 0$, and $\lambda_{22} = \mu \neq 0$, in the previous notation. Then the corresponding central simple algebra Q is called a (**generalized**) **quaternion algebra**. Thus Q has a basis $\{1, u, v, w\}$ satisfying the following laws: 1 is the unity element, $w = uv = -vu, u^2 = \lambda$, and $v^2 = \mu$ ($\lambda, \mu \in K$). Any central simple algebra of dimension 4 is isomorphic to a certain quaternion algebra. (In particular, if $K = \mathbf{R}$ and $\lambda = \mu = -1$, then Q coincides with the quaternion field \mathbf{H} .) For any element $x = \alpha + \beta u + \gamma v + \delta w$ of Q , the element $\bar{x} = \alpha - \beta u - \gamma v - \delta w$ is said to be **conjugate** to x , and $N(x) = x\bar{x} \in K$ is called the **norm** of x . An element x of Q is invertible in Q if and only if $N(x) \neq 0$.

E. Finite-Dimensional Associative Algebras over a Field

For the rest of this article, we assume that the algebras considered are unitary and finite

dimensional over a field K . Then by general properties of left (right) \dagger Artinian rings, any associative algebra A has the following structure: the \dagger radical N of A is the greatest \dagger nilpotent ideal, and the quotient algebra $A/N = \bar{A}$ is \dagger semisimple and decomposed into a direct sum of ideals which are simple algebras:

$$\bar{A} = \bar{A}_1 + \dots + \bar{A}_n.$$

Each simple component \bar{A}_i is a full matrix ring of degree r_i over a certain division algebra D_i , and \bar{A}_i is decomposed into a direct sum of r_i minimal left ideals which are mutually A -isomorphic:

$$\bar{A}_i = \bar{A}_i \bar{e}_i^{(1)} + \dots + \bar{A}_i \bar{e}_i^{(r_i)}, \quad 1 \leq i \leq n,$$

where $\bar{e}_i^{(1)}, \dots, \bar{e}_i^{(r_i)}$ are orthogonal \dagger idempotent elements of \bar{A}_i whose sum is equal to the unity element of \bar{A}_i . On the other hand,

$$\bar{A}_i = \bar{e}_i^{(1)} \bar{A} + \dots + \bar{e}_i^{(r_i)} \bar{A}$$

gives a decomposition of A_i into the direct sum of minimal right ideals that are mutually A -isomorphic. Moreover, we can choose an idempotent element $e_i^{(s)}$ of A from each residue class $\bar{e}_i^{(s)}$ such that $\{e_i^{(s)}\}$ forms a system of orthogonal idempotent elements, whose sum is equal to the unity element 1 of A , and

$$A = \sum_{i=1}^n \sum_{s=1}^{r_i} A e_i^{(s)} = \sum_{i=1}^n \sum_{s=1}^{r_i} e_i^{(s)} A$$

gives a decomposition of A into a direct sum of \dagger direct indecomposable left (right) ideals. Here, $A e_i^{(s)}$ and $A e_j^{(t)}$ ($e_i^{(s)} A$ and $e_j^{(t)} A$) are A -isomorphic if and only if $i = j$. Conversely, every decomposition of A into indecomposable ideals is obtained as above. The A -submodule $N e_i^{(s)}$ of $A e_i^{(s)}$ is the unique maximal proper submodule, and $A e_i^{(s)}/N e_i^{(s)}$ and $A e_j^{(t)}/N e_j^{(t)}$ are A -isomorphic if and only if $i = j$. Any simple A -module is A -isomorphic to a certain $A e_i^{(s)}/N e_i^{(s)}$ (- Sections H, I).

Any simple algebra A over K is isomorphic to a full matrix ring $M_n(D)$ over a certain division algebra D . This is called **Wedderburn's theorem**. Here n is determined uniquely by A , and D is also determined uniquely by A up to isomorphism. Moreover, the center of A is isomorphic to the center of D . If the center of A coincides with K , then A is called a **central simple algebra** (or **normal simple algebra**) over K . In this case, any isomorphism of two simple subalgebras of A can be extended to an \dagger inner automorphism of A . Let $V(B)$ denote the \dagger commutator of a simple subalgebra B of A . Then $V(B)$ is also a simple subalgebra and $V(V(B)) = B$, $\dim A = \dim B \cdot \dim V(B)$. In particular, if B is central over K , then there is a canonical isomorphism $A \cong B \otimes_K V(B)$. If D is a central division algebra, then any maximal commutative subalgebra L of D is a field and satisfies

$(\dim L)^2 = \dim D$. Moreover, there are \dagger separable extensions over K among such L . In general, the dimension of a central simple algebra A is a square number r^2 , where r is called the **degree** of A .

Two central simple algebras are said to be **similar** if they are isomorphic to full matrix rings over the same division algebra. This is an equivalence relation, and each equivalence class is called an **algebra class**. On the other hand, if A is a central simple algebra and B is a simple algebra, then $A \otimes_K B$ is a simple algebra. Moreover, if B is central, then $A \otimes_K B$ is also a central simple algebra. If A and B are similar to A' and B' , respectively, then $A \otimes_K B$ is similar to $A' \otimes_K B'$. Hence, the tensor product \otimes defines multiplication in the set $\mathcal{B}(K)$ of the algebra classes over K . If A is a central simple algebra and A° is an algebra anti-isomorphic to A , then A° is also central simple and $A \otimes_K A^\circ$ is isomorphic to a full matrix ring over K . This shows that $\mathcal{B}(K)$ forms a group, which is called the **Brauer group**, after R. Brauer, who introduced this concept, or the **algebra class group** over K . For a central simple algebra A , the degree of a division algebra similar to it is called the **Schur index** of A (or of the algebra class of A), and the order of the algebra class of A in the Brauer group is called the **exponent** of A . The Schur index is divisible by the exponent; and conversely, the exponent is divisible by every prime divisor of the Schur index.

F. Extensions of Coefficient Fields

Let L be an \dagger extension field over a field K . Then for any algebra A over K , $L \otimes_K A$ can be regarded as an algebra over L . This algebra is denoted by A^L and is called an **algebra obtained by extending the coefficient field** to L . Let us denote the radical of a ring A by $\mathfrak{R}(A)$. An algebra A over K is called a **separable algebra** if it satisfies $\mathfrak{R}(A^L) = \{0\}$ for any extension field L over K . In the special case when A is an \dagger algebraic extension field over K , A is a separable algebra if and only if every element of A (or every element of a subset which generates A) is \dagger separable over K (- 149 Fields). A (finite-dimensional) algebra A over K is separable if and only if A is semisimple and the center of every simple component of A is a separable extension over K . If the quotient algebra $A/\mathfrak{R}(A)$ of an algebra A is separable, then there exists a subalgebra S such that $A = S + \mathfrak{R}(A)$ and $S \cap \mathfrak{R}(A) = \{0\}$, and S is uniquely determined up to inner automorphisms (**Wedderburn-Mal'tsev theorem**).

In order that an algebra A over K be central simple, it is necessary and sufficient that A^L be

simple for any extension field L over K . This latter statement holds if and only if A^L is isomorphic to a full matrix algebra over L for a certain extension field L over K . Such an extension field is a splitting field of A (- 362 Representations F). For a central simple algebra A , a (finite) extension field L of degree r over K is a splitting field of A if and only if there exists a central simple algebra B of degree r which is similar to A having a subfield which is K -isomorphic to L . In this case, r is divisible by the Schur index of A . Furthermore, A has a separable splitting field whose extension degree is equal to the Schur index of A . This shows that A has a splitting field which is a finite Galois extension field over K .

Let L be a finite \dagger Galois extension over a field K and G be its \dagger Galois group. If f is a factor set with respect to the operation of G on L^* , then the crossed product $(L/K, f)$ is a central simple algebra over K (- Section D). The product fg of two factor sets f and g is also a factor set, and the set of all factor sets forms an \dagger Abelian group. Thus $(L/K, fg)$ is similar to $(L/K, f) \otimes_K (L/K, g)$. On the other hand, factor sets f and g are said to be **associated** with each other if there exists a family $\{\lambda_s\}_{s \in G}$ of elements of L such that

$$f(r, l) = g(r, l)r(\lambda_l)\lambda_{r^{-1}}^{-1}\lambda_r, \quad r, l \in G.$$

Hence, f and g are associated with each other if and only if $(L/K, f)$ and $(L/K, g)$ are similar. Therefore, the mapping $f \rightarrow (L/K, f)$ gives a monomorphism of the group $H^2(G, L^*)$ of all associated classes of factor sets (which can be identified with the 2-dimensional \dagger cohomology group of G with coefficients in L^*) into the Brauer group $\mathcal{B}(K)$ over K . Its image coincides with the subgroup of all algebra classes which have L as a splitting field. In particular, any algebra class is similar to the crossed product of a certain finite Galois extension L and its Galois group G (R. Brauer, E. Noether, A. Albert, K. Shoda, and others).

G. Cyclic Algebras

Let Z be a \dagger cyclic extension field of degree n over a field K . Then the crossed product of Z and its Galois group G is called a **cyclic algebra** over K . For a fixed generator s of G and any element $\alpha \neq 0$ of K , a factor set $f(s^i, s^j)$ ($0 \leq i, j < n$) can be defined by $f(s^i, s^j) = 1$ ($i+j < n$) and $f(s^i, s^j) = \alpha$ ($i+j \geq n$). Let (Z, s, α) denote the corresponding crossed product. Then (Z, s, α) and (Z, s, β) are similar if and only if α/β is a norm of a certain element of Z into K . On the other hand, any crossed product of Z and G is similar to a certain (Z, s, α) , and the correspondence $\alpha \rightarrow (Z, s, \alpha)$

gives an isomorphism of $K^*/N_{Z/K}(Z^*)$ (the norm class group of Z/K) to the group consisting of the algebra classes over K which have Z as a splitting field. If K is a $\dagger p$ -adic field or a finite \dagger algebraic number field, then any central simple algebra over K is isomorphic to a certain cyclic algebra. In this section, we describe this situation in detail.

Let K be a p -adic field and q the number of residue classes modulo p . If A is a central simple algebra over K , then its Schur index coincides with the exponent, which is called simply the **index**. A finite extension field L over K is a splitting field for A if and only if the degree of L is a multiple of the index of A . If n is the degree of A , then the field $W = K(w)$ obtained by adjoining to K a \dagger primitive (q -1)st root of unity ω is a cyclic (and \dagger unramified) extension of degree n over K , and its Galois group is generated by the automorphism σ of W determined by $\omega^\sigma = \omega^q$, $\sigma|_K = \text{identity}$. Moreover, we have $A \cong (W, \sigma, \alpha)$ for a certain $\alpha \in K^*$. Let v be the \dagger exponential p -adic valuation $v(\alpha)$ of α . Then $v/n \pmod{\mathbf{Z}}$ is uniquely determined by the algebra class of A , which is called the **Hasse invariant** of A (or of the algebra class of A). By assigning to each of the algebra classes its Hasse invariant, we get an isomorphism of the Brauer group $B(K)$ over K to the group \mathbf{Q}/\mathbf{Z} , the additive group of the rational numbers mod \mathbf{Z} (H. Hasse, 1931).

Let K be a finite algebraic number field and A be a central simple algebra over K . Let \mathfrak{p} be a (finite or infinite) \dagger prime divisor of K and $K_{\mathfrak{p}}$ the $\dagger p$ -adic extension field over K . The algebra A , which is obtained from A by extending the coefficient field to $K_{\mathfrak{p}}$ is a central simple algebra over K . Except for a finite number of \mathfrak{p} , A is isomorphic to a full matrix ring over $K_{\mathfrak{p}}$, and A itself is isomorphic to a full matrix ring over K if and only if A is isomorphic to a full matrix ring over $K_{\mathfrak{p}}$ for all \mathfrak{p} . The index $m_{\mathfrak{p}}$ of A is called the \mathfrak{p} -index of A , and the Hasse invariant of A is called the **\mathfrak{p} -invariant** of A , which is denoted by (A/\mathfrak{p}) . If \mathfrak{p} is an infinite prime divisor, then $m_{\mathfrak{p}}$ is equal to 1 or 2, and in each case, we define the \mathfrak{p} -invariant by setting $(A/\mathfrak{p}) = 0$ or $1/2 \pmod{\mathbf{Z}}$ correspondingly. The Schur index of A is the L.C.M. of the \mathfrak{p} -indices $m_{\mathfrak{p}}$ for all \mathfrak{p} and coincides with the exponent of A . This is called simply the **index** of A . On the other hand, the \mathfrak{p} -invariants satisfy $(A/\mathfrak{p}) \equiv 0 \pmod{\mathbf{Z}}$ except for a finite number of \mathfrak{p} , and

$$\sum_{\mathfrak{p}} (A/\mathfrak{p}) \equiv 0 \pmod{\mathbf{Z}}.$$

Conversely, given a rational number $\rho_{\mathfrak{p}}$ for each \mathfrak{p} such that (i) $\rho_{\mathfrak{p}} \equiv 0 \pmod{\mathbf{Z}}$ except for a finite number of \mathfrak{p} ; (ii) $\rho_{\mathfrak{p}} \equiv 0 \pmod{\mathbf{Z}}$ if \mathfrak{p} is infinite and imaginary, $\rho_{\mathfrak{p}} \equiv 0$ or $1/2 \pmod{\mathbf{Z}}$ if

p is infinite and real; (iii) $\sum_p \rho_p \equiv 0 \pmod{\mathbf{Z}}$, then there is a uniquely determined algebra class of central simple algebras A over K such that $(A/p) \equiv \rho_p \pmod{\mathbf{Z}}$ for each p . In this way the structure of the Brauer group over a finite algebraic number field is completely determined (Hasse, 1933).

H. Frobenius Algebras

Let A be an algebra over a field K . A is called a **Frobenius algebra** if its \dagger regular representation and \dagger coregular representation (\rightarrow 362 Representations E) are similar. Thus A is a Frobenius algebra if the left A -module A and the dual module A^* of the right A -module A are isomorphic as left A -modules. Let

$$A = \sum_{i=1}^n \sum_{s=1}^{r_i} Ae_i^{(s)} = \sum_{i=1}^n \sum_{s=1}^{r_i} e_i^{(s)} A$$

be direct decompositions of A into indecomposable left (right) ideals (\rightarrow Section E). We denote $e_i^{(1)}$ by e_i . Then A is a Frobenius algebra if and only if there exists a permutation π on $1, \dots, n$ such that (i) $Ae_i \cong (e_{\pi(i)}A)^*$; (ii) $r_i = r_{\pi(i)}$. When there exists a permutation satisfying only condition (i), A is called a **quasi-Frobenius algebra**.

For a subset S of A , the ideals $I(S) = \{a \in A \mid aS = 0\}$ and $r(S) = \{a \in A \mid Sa = 0\}$ are called the **left annihilator** and **right annihilator** of S , respectively. Then A is a quasi-Frobenius algebra if and only if (ii) $l(r(I)) = I$ and $r(l(r)) = r$ for any left ideal I and any right ideal r . In general, if we are given a left (right) A -module M , we denote the right (left) A -module $\text{Hom}_A(M, A)$ by f_i . Then if A is a quasi-Frobenius algebra, there is a canonical isomorphism $\hat{M} \cong M$, and the annihilator relation gives a one-to-one and dual correspondence between the set of the submodules of M and the set of submodules of \hat{M} (M. Hall). If a quasi-Frobenius algebra A satisfies (iv) $\dim r + \dim I(r) = \dim l + \dim r(l) = \dim A$ for any left ideal I and any right ideal r , then A is a Frobenius algebra; the converse is also true.

A criterion for an algebra A to be a Frobenius algebra is that there is a linear form $x \rightarrow \lambda(x)$ on A such that if $\lambda(xa) = 0$ for all $x \in A$, then $a = 0$. Moreover, if λ satisfies $\lambda(xy) = \lambda(yx)$ ($x, y \in A$), then A is called a **symmetric algebra**. For example, semisimple algebras and group algebras are symmetric algebras. If A is a symmetric algebra, then for any left (right) A -module M the right (left) A -modules $M^* = \text{Hom}_K(M, K)$ and $\hat{M} = \text{Hom}_A(M, A)$ are canonically A -isomorphic.

If A is a Frobenius algebra, the radical N of A satisfies $l(N) = r(N)$, and the annihilator of N is a principal left and principal right ideal;

the converse is also true. For a two-sided ideal z of a Frobenius algebra A , the quotient algebra A/z is a Frobenius algebra if and only if $l(z)$ and $r(z)$ are a principal left and a principal right ideal, respectively. If A is a symmetric algebra, then any two-sided ideal z satisfies $l(z) = r(z)$, and A/z is also a symmetric algebra if and only if $l(z) = r(z)$ is a principal ideal generated by an element in the center.

Furthermore, for any extension L of the coefficient field K , A^L is a quasi-Frobenius algebra (resp. Frobenius algebra, symmetric algebra) if and only if A is (T. Nakayama, 1939, 1941). The concept of Frobenius algebras has been extended to algebras B over a ring A (F. Kasch, 1954).

I. Uniserial Algebras

In the notation of the preceding section, if each indecomposable left ideal Ae_i (right ideal e_iA) of an algebra A over K has a unique composition series, then A is called a **generalized uniserial algebra**. If an algebra A is decomposed into a direct sum of ideals which are primary rings, then it is called a **uniserial algebra**. Any left module over a generalized uniserial algebra A is decomposed into a direct sum of submodules which are A -homomorphic images of Ae_i . An algebra whose radical N is a principal left and principal right ideal is a generalized uniserial algebra. For an algebra A to be uniserial, it is necessary and sufficient that every two-sided ideal of A be a principal left and principal right ideal. Hence A is uniserial if and only if every quotient algebra of A is a Frobenius algebra. If A^L is uniserial for an extension field L of K , then A itself is uniserial. The converse, however, is not always true. If A^L is uniserial for any extension field L of K , then A is called an **absolutely uniserial algebra**. For A to be absolutely uniserial it is necessary and sufficient that its radical N be a principal ideal generated by an element in the center Z and Z be decomposed into a direct sum of simple extensions of K (i.e., ideals of the form $K[a]$) (K. Asano, G. Köthe, Nakayama, G. Azumaya).

J. Algebraic Algebras

Here, we consider general (not necessarily finite-dimensional) algebras A over a field K . We say that A is an **algebraic algebra** if every element of A is algebraic over K , i.e., every element of A is a root of a certain polynomial with coefficients in K . We say that A satisfies a **polynomial identity** $p(X_1, \dots, X_n) = 0$ or that A is a **PI-algebra** with an identity $p(X_1, \dots, X_n) =$

0 if there exists a nonzero (noncommutative) polynomial $p(X_1, \dots, X_n)$ in X_1, \dots, X_n with coefficients in K such that $p(a_1, \dots, a_n) = 0$ for all $a_i \in A$. A PI-algebra satisfies an identity which is homogeneous linear in each variable, and also an identity of the form $[x_1, \dots, x_n]^m = 0$ (where $[\]$ is the sum $\sum \pm x_{i_1} \dots x_{i_n}$ taken over all permutations of $1, 2, \dots, n$ and \pm is the sign of the permutation). An algebra is said to be **locally finite** if any finite number of elements of A generate a finite-dimensional subalgebra. For PI-algebras, an affirmative answer was found for Kurosh's problem, which asks whether an algebraic algebra is locally finite if the degree of any element α of A (i.e., $\dim K[a]$) is bounded.

K. Brauer Group of a Ring

Let R be a commutative ring. An R -algebra is called a **separable algebra** if A is \dagger projective as a two-sided A -module (- 200 Homological Algebra F). When the base ring is a field, this agrees with the classical notion of separability; and A is separable over R if and only if $A/\mathfrak{m}A$ is separable over the residue field R/\mathfrak{m} for every maximal ideal \mathfrak{m} of R . A **central separable algebra** is also called an **Azumaya algebra**. If P is a finitely generated faithful projective R -module (briefly: **R-progenerator**), the endomorphism ring $\text{End}_R(P)$ is an Azumaya R -algebra. Azumaya algebras A , and A , are said to be in the same class (similar) if there exist R -progenerators P_1 and P_2 such that $A \otimes \text{End}_R(P_1) \cong A \otimes \text{End}_R(P_2)$. The set of similarity classes forms an Abelian group with respect to \otimes . This is called the **Brauer group** $B(R)$ of R (Auslander and Goldman [10]). Every element of $B(R)$ is of finite order [12, 13].

$B(R)$ is a covariant functor from commutative rings to Abelian groups. If R is a field, $B(R)$ coincides with the classically defined one (- Section E). If R is a \dagger Henselian local ring with the residue field k , the mapping $B(R) \rightarrow B(k)$ is an isomorphism [9]. If R is a \dagger regular ring with the quotient field K , $B(R) \rightarrow B(K)$ is injective. If further $\dim R \leq 2$, we have $B(R) = \bigcap_{\mathfrak{p}} B(R_{\mathfrak{p}})$, \mathfrak{p} running over all primes of height 1 of R , where $R_{\mathfrak{p}}$ is the localization of R at \mathfrak{p} and $B(R)$ and $B(R_{\mathfrak{p}})$ are considered as embedded in $B(K)$ [10]. As an example, putting these facts together with the structure of $B(K)$ of the algebraic number field K (\rightarrow Section G), we have the structure of $B(R)$ of the ring of integers of K : $B(R) = 0$ if K is totally imaginary, and $\cong (\mathbb{Z}/2\mathbb{Z})^{r-1}$ if K has $r(>0)$ real infinite places.

A commutative R -algebra S is called a **splitting ring** of A if the S -algebra $S \otimes A$ is isomor-

phic to $\text{End}_S(P)$ for some S -progenerator P . We denote by $B(S/R)$ the subgroup of $B(R)$ consisting of all algebra classes split by S . Since an Azumaya algebra over a ring need not have a Galois extension as splitting ring, the description of the Brauer groups by means of Galois cohomology ceases to have full generality. Instead we have the following exact sequence of \dagger Amitsur cohomology, assuming S is an R -progenerator (Chase and Rosenberg [11]): $0 \rightarrow H^1(S/R, U) \rightarrow \text{Pic}(R) \rightarrow H^0(S/R, \text{Pic}) \rightarrow H^2(S/R, U) \rightarrow B(S/R) \rightarrow H^1(S/R, \text{Pic}) \rightarrow H^3(S/R, U) \rightarrow$, where we denote by $U(\dagger')$ and $\text{Pic}(T)$ of a commutative ring T the unit group and the \dagger Picard group of rank 1 projective modules of T , respectively. The full Brauer group $B(R)$ is mapped monomorphically into $H^2(R, U) = \varinjlim H^2(S/R, U)$, the limit over the \dagger faithfully flat R -algebras S .

Grothendieck and others studied the Brauer groups in a more general geometrical context [12].

References

- [1] A. A. Albert, Structure of algebras, Amer. Math. Soc. Colloq. Publ., 1939.
- [2] E. Artin, C. J. Nesbitt, and R. Thrall, Rings with minimum condition, Univ. of Michigan Press, 1944.
- [3] M. Deuring, Algebren, Erg. Math., Springer, second edition, 1968 (Chelsea, 1948).
- [4] N. Jacobson, The theory of rings, Amer. Math. Soc. Math. Surveys, 1943.
- [5] N. Jacobson, Structure of rings, Amer. Math. Soc. Colloq. Publ., 1956.
- [6] B. L. van der Waerden, Algebra II, Springer, 1967.
- [7] C. W. Curtis and I. Reiner, Representation theory of finite groups and associative algebras, Interscience, 1962.
- [8] J.-P. Serre, Corps locaux, Actualités Sci. Ind., Hermann, 1962.
- [9] G. Azumaya, On maximally central algebras, Nagoya Math. J., 2 (1951), 119–150.
- [10] M. Auslander and O. Goldman, the Brauer group of a commutative ring, Trans. Amer. Math. Soc., 97 (1960), 367–409.
- [11] S. Chase and A. Rosenberg, Amitsur cohomology and the Brauer group, Mem. Amer. Math. Soc., 52 (1965), 34–79.
- [12] A. Grothendieck, Le groupe de Brauer, I, II, III, Dix exposés sur la cohomologie des schémas, North-Holland, 1968, 46–188.
- [13] M.-A. Knus and M. Ojanguren, Théorie de la descente et algèbres d'Azumaya, Lecture notes in math. 389, Springer, 1974.
- [14] M. Orzech and C. Small, the Brauer group of commutative rings, Lecture notes in pure and appl. math. 11, Dekker, 1975.

30 (X.19) Asymptotic Series

A. Asymptotic Series and Asymptotic Expansions

Let $\varphi_n(x)$, $n=0, 1, 2$, be functions defined in \mathbf{R}_+ . The sequence of functions $\{\varphi_n(x)\}_{n=0}^\infty$ is called an **asymptotic sequence** for $x \rightarrow \infty$ if

$$\varphi_{n+1}(x) = o(\varphi_n(x)) \text{ as } x \rightarrow \infty \quad (1)$$

for each n . When $\{\varphi_n(x)\}_{n=0}^\infty$ is an asymptotic sequence, the formal series $\sum a_n \varphi_n(x)$ is called an **asymptotic series**. In most cases, the functions $\varphi_n(x)$ have the form $\psi(x)\varphi(x)^n$. When $\varphi_n(x) = x^{-n}$ an asymptotic series is called an **asymptotic power series**. A function $f(x)$ defined on \mathbf{R}_+ is said to have the **asymptotic expansion** (or be **asymptotically developable** in the form)

$$f(x) \sim a_0 \varphi_0(x) + a_1 \varphi_1(x) + \dots + a_n \varphi_n(x) + \dots \quad (2)$$

as $x \rightarrow \infty$ if $f(x)$ satisfies

$$\begin{aligned} f(x) - a_0 \varphi_0(x) - a_1 \varphi_1(x) - \dots - a_n \varphi_n(x) \\ = O(\varphi_{n+1}(x)) \end{aligned} \quad (3)$$

for any integer $n \geq 0$ as $x \rightarrow \infty$. The coefficients a_n ($n = 0, 1, \dots$) appearing in (2) are uniquely determined. This fact immediately follows from the formulas

$$a_0 = \lim_{x \rightarrow \infty} f(x)/\varphi_0(x), \dots,$$

$$\begin{aligned} a_n = \lim_{x \rightarrow \infty} (f(x) - a_0 \varphi_0(x) - \dots \\ - a_{n-1} \varphi_{n-1}(x))/\varphi_n(x). \end{aligned}$$

For example, the † Hankel function of the first kind $H_v^{(1)}(x)$ has an asymptotic expansion

$$\begin{aligned} H_v^{(1)}(x) \\ \sim \frac{1}{\sqrt{\pi x}} \exp \left[i \left(x - \frac{\pi v}{2} - \frac{\pi}{4} \right) \right] \sum_{m=0}^{\infty} \frac{(v, m)}{(-2ix)^m}, \quad (4) \\ (v, m) = \frac{(4v^2 - 1)(4v^2 - 3^2) \dots [4v^2 - (2m - 1)^2]}{2^{2m} m!} \end{aligned}$$

If $v + \frac{1}{2}$ is not a positive integer, $\sum_{m=0}^{\infty} (v, m) (-2ix)^{-m}$ is not convergent, and (4) means that for each n ,

$$\begin{aligned} H_v^{(1)}(x) - \frac{1}{\sqrt{\pi x}} \exp \left[i \left(x - \frac{v\pi}{2} - \frac{\pi}{4} \right) \right] \sum_{m=0}^{n-1} \frac{(v, m)}{(-2ix)^m} \\ = O(x^{-n-1/2}) \end{aligned}$$

for $x \rightarrow \infty$.

Assume that $f(x)$ and $g(x)$ admit asymptotic expansions in power series for $x \rightarrow \infty$. Then all functions defined by $f(x) + g(x)$, $f(x) - g(x)$, $f(x)g(x)$, $f(x)/g(x)$ ($\lim_{x \rightarrow \infty} g(x) \neq 0$) admit

asymptotic expansions in power series for $x \rightarrow \infty$. Furthermore, if $f'(x)$ has an asymptotic expansion for $x \rightarrow \infty$, its expansion is obtained by termwise differentiation of the $f(x)$'s.

For any asymptotic power series $\sum_{n=0}^{\infty} a_n x^{-n}$ there always exists a smooth function $f(x)$ defined in \mathbf{R}_+ such that

$$f(x) \sim a_0 + a_1 x^{-1} + a_2 x^{-2} + \dots$$

for $x \rightarrow \infty$, and all the derivatives of $f(x)$ are asymptotically developable in power series. Indeed, a function defined by

$$f(x) = \sum_{n=0}^{\infty} a_n \theta(x/t_n) x^{-n}$$

satisfies the required properties if $\theta(t)$ is a smooth function satisfying

$$\theta(t) = \begin{cases} 0, & t \leq 1, \\ 1, & t \geq 2, \end{cases}$$

and the sequence $\{t_n\}_{n=0}^\infty$ satisfies $1 \leq t_0 \leq t_1 \leq \dots$ and

$$|n(n+1) \dots (2n-1) 2na_n t_n^{-1}| \leq 2^{-n}$$

for all n .

Until now only asymptotic series for $x \rightarrow \infty$ have been considered, but the notion of asymptotic series has been adapted in various forms to the spaces of functions under consideration.

(a) Asymptotic series in a complex domain. Let α be a boundary point of an open connected domain D in the complex z -plane. When the asymptotic behavior for $z \rightarrow \alpha$ of holomorphic functions in D is considered, we adopt as an asymptotic sequence $\{\varphi_n(z)\}_{n=0}^\infty$ of holomorphic functions in D satisfying for each n

$$\varphi_{n+1}(z) = o(\varphi_n(z))$$

as z tends to α through D . Let D be an angular domain with vertex at α . For any asymptotic power series $\sum_{n=0}^{\infty} a_n (z - \alpha)^n$ there always exists a function $f(z)$ that is holomorphic in D and that admits an asymptotic expansion

$$f(z) \sim a_0 + a_1 (z - \alpha) + \dots + a_n (z - \alpha)^n + \dots \quad (5)$$

for $z \rightarrow \alpha$. Concerning the uniform convergence of asymptotic expansions, **Carleman's theorem** is well known [2]: If the asymptotic expansion (5) is valid as z tends to α through an arbitrary angular domain in the Riemann surface of $\log(z - \alpha)$ having α as its vertex, then the asymptotic expansion (5) is uniformly convergent.

(b) Asymptotic expansions by a large parameter. In the theory of differential equations, functions with a large parameter are often considered. Let Ω be a domain in \mathbf{R}^n and k a parameter ≥ 1 , and let $\varphi_n(x, k)$, $n=0, 1, 2, \dots$, be functions defined for $(x, k) \in \Omega \times [1, \infty)$. If

they satisfy for each $x \in \Omega$ and n

$$\varphi_{n+1}(x, k) = o(\varphi_n(x, k)) \quad (6)$$

for $k \rightarrow \infty$, the sequence $\{\varphi_n(x, k)\}_{n=0}^{\infty}$ is called an asymptotic sequence, and the formal series $\sum_{n=0}^{\infty} a_n(x) \varphi_n(x, k)$ is called an asymptotic series. A function $f(x, k)$ is said to be asymptotically developable in the form

$$f(x, k) \sim a_0(x) \varphi_0(x, k) + a_1(x) \varphi_1(x, k) + \dots + a_n(x) \varphi_n(x, k) + \dots$$

for $k \rightarrow \infty$ if for every $x \in \Omega$ and every n

$$f(x, k) - a_0(x) \varphi_0(x, k) - \dots - a_{n-1}(x) \varphi_{n-1}(x, k) = O(\varphi_n(x, k)) \quad (7)$$

holds as $k \rightarrow \infty$. Especially, if the order relations (6) and (7) hold for each n uniformly in $x \in \Omega$, we say that $f(x, k)$ is uniformly asymptotically developable. In most cases, the $\varphi_n(x, k)$ are of the form $\psi(x, k) k^{-n}$.

B. Methods of Asymptotic Expansion

There are several general methods of obtaining asymptotic expansions of functions represented by integrals with a large parameter. These methods are useful in applications because most special functions are represented by integrals (- 39 Bessel Functions, 389 Special Functions).

(a) **Laplace's method.** Let $h(s)$ be a smooth real-valued function in $[a - \delta, a + \delta]$ ($\delta > 0$) satisfying

$$h(a) < h(s) \text{ for all } s \in [a - \delta, a + \delta] - \{a\}$$

and

$$h''(a) > 0,$$

and let $g(s) \in C^{\infty}([a - \delta, a + \delta])$. Then

$$I(x) = \int_{a-\delta}^{a+\delta} e^{-xh(s)} g(s) ds$$

has an asymptotic expansion

$$I(x) \sim e^{-xh(a)} \frac{\sqrt{2\pi}}{\sqrt{x}} \left\{ g(a) \frac{1}{\sqrt{h''(a)}} + c_1 x^{-1} + \dots + c_n x^{-n} + \dots \right\}$$

for $x \rightarrow \infty$, where the c_n are determined by derivatives of $h(s)$ at $s = a$ of order $\leq 2n + 2$ and those of $g(s)$ at $s = a$ of order $\leq 2n$.

As an application of this method, consider the Γ -function $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$. By a change of variable $t = xs$, we have $\Gamma(x) = x^x \int_0^{\infty} e^{-x(s-\log s)} s^{-1} ds$. Since for any $\delta > 0$

$$\left| \int_0^{1-\delta} e^{-x(s-\log s)} s^{-1} ds \right| + \left| \int_{1+\delta}^{\infty} e^{-x(s-\log s)} s^{-1} ds \right| \leq C e^{-x(1+\alpha)} \quad (\alpha > 0),$$

the asymptotic expansion by the Laplace method

$$\int_{1-\delta}^{1+\delta} e^{-x(s-\log s)} s^{-1} ds \sim e^{-x} \sqrt{\frac{2\pi}{x}} (1 + c_1 x^{-1} + \dots)$$

implies †Stirling's formula

$$\Gamma(x) = \sqrt{2\pi} x^{x-1/2} e^{-x} (1 + O(x^{-1})).$$

(b) **Stationary phase method.** Let $h(s)$ be a real-valued smooth function defined in $[a - \delta, a + \delta]$ such that

$$h'(u) = 0, h''(u) \neq 0 \text{ and } h'(s) \neq 0 \text{ for } s \neq a,$$

and let $g(s)$ be a function in $C_0^{\infty}(a - \delta, a + \delta)$. Then

$$I(x) = \int_{a-\delta}^{a+\delta} e^{ixh(s)} g(s) ds$$

has an asymptotic expansion

$$I(x) \sim e^{ixh(a)} e^{i\sigma\pi/4} \frac{1}{\sqrt{x}} \left\{ \frac{g(a)}{\sqrt{|h''(a)|}} + c_1 x^{-1} + \dots + c_n x^{-n} + \dots \right\},$$

where $\sigma = h''(a)/|h''(a)|$ and the c_n are determined by derivatives of $h(s)$ at $s = a$ of order $\leq 2n + 2$ and those of $g(s)$ at $s = a$ of order $\leq 2n$.

Further, †integration by parts and the †method of steepest descent are often used to derive asymptotic expansions of integrals with a large parameter.

C. Application to Differential Equations

For a system of linear ordinary differential equations with an †irregular singular point at $z = 0$, even when there exists a formal power series solution, the power series is generally divergent. Taking the hint given by Stirling's formula for the Γ -function, Poincaré introduced the notion of asymptotic expansions and succeeded in giving an analytic meaning to formal solutions of divergent type. Actual solutions for †difference equations, †difference-differential equations, and †ordinary differential equations of canonical form (including second-order linear †differential equations of confluent type) always have asymptotic expansions.

In the study of certain ordinary differential equations containing a large parameter k , an asymptotic solution is often obtained in the form

$$u(x, k) \sim e^{ik\varphi(x)} (a_0(x) + a_1(x)k^{-1} + \dots + a_n(x)k^{-n} + \dots)$$

(†WKB method). Also, in the study of linear

partial differential equations, an asymptotic solution of the above form plays an important role. For example, consider the Schrodinger equation

$$L_\lambda[u] = \frac{1}{i\lambda} \frac{\partial u}{\partial t} - \frac{1}{(i\lambda)^2} \Delta u - v(x)u = 0.$$

An asymptotic series

$$u(x, t; \lambda) \sim e^{i\lambda\phi(x, t)} \left(a_0(x, t) + a_1(x, t) \frac{1}{i\lambda} + \dots + a_n(x, t) \left(\frac{1}{i\lambda} \right)^n + \dots \right) \quad (8)$$

satisfies $L_\lambda[u] \rightarrow 0$ for $\lambda \rightarrow \infty$ if

$$\frac{\partial \phi}{\partial t} (\nabla \phi)^2 - v(x) = 0 \quad (\text{eikonal equation}),$$

$$\frac{\partial a_n}{\partial t} - 2\nabla \phi \cdot \nabla a_n - \Delta \phi a_n = \Delta a_{n-1}, \quad n=0, 1, 2, \dots \quad (\text{transport equation}),$$

where we set $a_{-1} \equiv 0$. For each asymptotic solution of the form (8) there exists an actual solution of $L_\lambda[f(x, t; \lambda)] = 0$ which is asymptotically developable in the form

$$f(x, t; \lambda) \sim \sum_{n=0}^{\infty} e^{i\lambda\phi(x, t)} a_n(x, t) (i\lambda)^{-n}$$

Then asymptotic solutions of the form (8) can describe the fundamental properties of the phenomenon governed by the equation $L_\lambda[u] = 0$.

†Geometric optics can also be illustrated by asymptotic solutions of the form (8) (\rightarrow 325 Partial Differential Equations of Hyperbolic Type L). The ideas of asymptotic series, asymptotic expansions, and asymptotic solutions are essential to the theory of †pseudodifferential operators and †Fourier integral operators (\rightarrow 345 Pseudodifferential Operators).

References

- [1] A. Erdélyi, *Asymptotic expansions*, Dover, 1956.
- [2] T. Carleman, *Les fonctions quasi-analytiques*, Gauthier-Villars, 1926.
- [3] H. Poincaré, *Sur les intégrales irrégulières des équations linéaires*, *Acta Math.*, 8 (1886), 295–344.
- [4] W. R. Wasow, *Asymptotic expansions for ordinary differential equations*, Interscience, 1965.
- [5] W. B. Ford, *The asymptotic developments of functions defined by Maclaurin series*, Scientific Series, Univ. of Michigan, 1936 (Chelsea, 1960).
- [6] M. Hukuhara, T. Kimura, and T. Matuda, *Equations différentielles ordinaires du premier*

ordre dans le champ complexe, *Publ. Math. Soc. Japan*, 1961.

[7] F. W. J. Olver, *Asymptotic and special functions*, Academic Press, 1974.

[8] G. D. Birkhoff, *Collected mathematical papers*, Dover, 1968.

[9] V. P. Maslov, *Théorie des perturbations et méthodes asymptotiques*, Dunord, Gauthier-Villars, 1972. (Original in Russian, 1965.)

31 (XVI.1) Automata

A. General Remarks

Automaton is a generic term for the mathematical models for automatic machines having memory devices and capable of performing definite acts within each unit of time. There are several types of automata, namely, Turing machines for the theory of computability [6], linear bounded automata, push-down automata, and finite automata for mathematical language theory [10–13]. These models are also used in the theory of †complexity of computation of algorithms [14].

B. Turing Machines

Since the arithmetization of metamathematics was used in the proof of †Gödel's incompleteness theorem, mathematical formulations of the notion of effectively computable functions have been attempted by many mathematicians, including S. C. Kleene, A. Church, and others, as well as by Gödel himself (\rightarrow 356 Recursive Functions). With this end in mind, A. M. Turing and E. L. Post independently introduced a sort of ideal computer [1, 2], called a **Turing machine**.

Such a machine is capable of being in one of a finite number of **internal states** (possible contents of a memory device) q_0, q_1, \dots, q_n fixed for the machine; and the machine is supplied with a tape divided into squares. The tape is infinite in both directions. Each square can be blank or can have printed on it any one of the previously specified symbols s_0, s_1, \dots, s_m , where s_0 stands for the blank. In any situation the number of nonblank squares is finite. Let $Q = \{q_0, q_1, \dots, q_n\}$, $M = \{s_0, s_1, \dots, s_m\}$, and consider the mapping $t: Q \times M \rightarrow Q \times M \times \{R, L, 0\}$. A Turing machine T is specified by (Q, M, t, q_0, F) , where F is a subset of Q called a set of final states and q_0 is the specified initial state. If $t(q, s) = (q', s', x)$, then the 5-tuple (q, s, q', s', x) is called an **instruction** for T . The

machine acts as follows: (1) At the start, the machine is in the state q_0 . Some sequence of symbols in M are written on the tape as an input, and the machine is on the square next to the leftmost nonblank symbol. (2) With respect to the present state q and the scanned symbol, the machine selects an instruction (q, s, q', x) ; then the next state becomes q' , s is replaced by s' , and the machine moves one square right or left or stands still according as x is R , L , or 0 . (3) The machine stops when the state becomes an element of F , and then the sequence of symbols on the tape is the result of the computation and considered to be the output. A Turing machine is called **deterministic** if there is always at most one instruction associated with a given pair (q, s) ; otherwise it is called **nondeterministic**.

We represent a natural number y on the tape by $y + 1$ squares, each of which has the symbol s_1 printed on it. A pair (y_1, y_2) is represented on the tape by the following sequence of symbols of length $y_1 + y_2 + 5$: one s_0 , the representation of y_1 ; one s_0 , the representation of y_2 ; and one s_0 . Similarly, in general we can represent a k -tuple (y_1, \dots, y_k) of natural numbers on the tape. Let Φ be a partial function of k variables, $k \geq 0$. We say that a Turing machine T **computes** Φ if T starts with the tape in which the representation of (x_1, \dots, x_k) is printed and stops with the tape having the representation of (x_1, \dots, x_k, x) when $\Phi(x_1, \dots, x_k)$ is defined and equal to x . A partial function is **computable** (in the sense of Turing) if there exists a Turing machine that computes the function. It is known that a partial function is computable if and only if it is partial recursive. The equivalence of computability (in the sense of Turing) and recursiveness for number-theoretic functions is strong evidence for the validity of Church's thesis (- 356 Recursive Functions).

We can apply a Turing machine not only to the computation of number-theoretic functions but also to the transformation of a finite sequence of letters in any language with a finite list of letters. A finite nonempty list of distinct symbols is called an **alphabet**, and a finite sequence of symbols in the alphabet is called a **word** in that alphabet. All the words in an alphabet constitute a 'free monoid if we add the empty sequence and think of the concatenation of two words as its multiplication. We call a finite nonempty list $(A_1, B_1), \dots, (A_n, B_n)$ of pairs of words a set of **production rules**. By an **application** of a rule (A_i, B_i) to a word W , we mean that an occurrence of A_i exists in W , and by the application, A_i is replaced by B_i . We say that two words R and S are **equivalent** if they are transformable into each other by a finite number of applications of the production

rules. The following problem is called **Thue's (general) problem**: Is there an algorithm for deciding, for any given alphabet and a set of production rules, whether or not any given two words are equivalent? E. L. Post and A. A. Markov proved by applying Turing machines that this decision problem is unsolvable [4, 3] (a negative solution for the so-called word problem for semigroups).

C. Universal Turing Machine and Generalizations

Turing showed that it is possible to construct a machine U with the following property [1]: If we supply U with a tape on which a suitably coded list of instructions of any given machine T is printed then U successively prints on it the behavior of T which results from the execution of the instructions of T . In other words, U simulates the operations performed by any machine. Such a machine U is called a **universal Turing machine**. We can construct a universal Turing machine with only two internal states and also one having only two symbols, including the blank s_0 (C. E. Shannon and J. McCarthy [7]).

Let ψ_1, \dots, ψ_k be given functions that are all defined everywhere on the domain considered. Then we can relativize (\rightarrow 356 Recursive Functions) the notion of computability of functions (in the sense of Turing) to ψ_1, \dots, ψ_k . Indeed, we can consider a machine acting as follows: When the machine is in a state q_{ij} ($j = 1, \dots, k$ and we assume that these states are specified correspondingly to ψ_1, \dots, ψ_k and are not in F), then the machine prints the representation of $\psi_j(y_1, \dots, y_n)$ to the right of (y_1, \dots, y_n) and scans the resulting $(n_j + 1)$ -tuple $(y_1, \dots, y_n, \psi_j(y_1, \dots, y_n))$. In this process, the symbol a that is written to the right of (y_1, \dots, y_n) is moved one step to the right so that a is preserved on the right of the resulting $(n + 1)$ -tuple. Extending the preceding, Kleene defined a machine that computes a functional of finite types [S]. A functional of variables of any finite types is computable by such a generalized Turing machine if and only if it is partial recursive.

D. Language and Automata

Let M and N be mutually disjoint alphabets and $V = MU^*N$. Then $G = (N, M, P, S)$ is called a **Chomsky grammar** if P is a set of production rules consisting of finite number of pairs of elements of V^* and S is specified element of N , where V^* means the free monoid generated by V . For a word $u \in V^*$, if the result of a finite

Automata

number of applications of rules in P is v , then we write $u \rightarrow v$, and v is said to be **derived** from u by G . The **language** generated by G is $L(G) = \{u \mid S \rightarrow u \text{ and } u \in M^*\}$. N. Chomsky classified the grammars into three families with respect to the form of production rules: (1) **Context-sensitive grammar**: For each $(u, v) \in P$, the length of u is less than or equal to the length of v . (2) **Context-free grammar**: For each $(u, v) \in P$, u is an element of N and v is not a empty word. (3) **Regular grammar**: For each $(u, v) \in P$, $u \in N$ and v is either an element of M or of the form ax , where $a \in M$ and $x \in N$. Let $F_1(M)$, $F_2(M)$, and $F_3(M)$ be the families of languages generated by the grammars in the corresponding families classified as above. Then $F_1(M) \supsetneq F_2(M) \supsetneq F_3(M)$ holds and this is called a **Chomsky hierarchy** [10]. Formal language theory studies various structures of these families of languages.

Corresponding to the three families of languages, we have the following families of automata: (1) A **deterministic (nondeterministic) linear bounded automaton** is a Turing machine with input tape, namely, $T = (Q, M \cup N, t, q_0, F)$, where Q is the set of states, M an input alphabet, N a tape alphabet, t is a mapping from $Q \times (M \cup \{e\}) \times (N \cup \{e\})$ into $Q \times N \times \{L, R\}$ (the power set of $Q \times N \times \{L, R\}$), $q_0 \in Q$ is the initial state, and F is a subset of Q . A word $w \in M^*$ is placed on the input tape, and T starts its computation from the initial state with blank tape and by reading one symbol from the input tape, and with respect to the mapping t , we decide the configuration of the next step. If at a certain stage, T reaches a state in F when T reads up w and if the length of the tape used is a constant multiple of the length of the input tape, then w is said to be **accepted** by T . (The special symbol e makes it possible to change the configuration without moving the input tape.) (2) A **push-down automaton** is a Turing machine with an input tape. The working tape is of infinite length and moves in only one direction, say to the right. The computation with input word $w \in M^*$ starts with the position at the left end of the work tape. It uses the tape according to a first-in-first-out principle; namely, the machine may write a symbol to the blank square, but if it is on the rightmost nonblank square, it may not go left without erasing the nonblank symbol. If it reaches a state in F , then w is said to be **accepted** by T . (3) A **finite automaton** is a machine without its infinite tape; a word w is **accepted** if T reaches a state in F (without using the symbol e).

Let $L(T)$ be the set of words over M accepted by T ; then $L(T)$ is called the **language accepted** by T . Corresponding to the three families of automata, $DF_i(M)$ and $NF_i(M)$, $i =$

1, 2, 3, denote the family of languages accepted by deterministic or nondeterministic automata. It is known that $DF_i(M) = NF_i(M) = F_i(M)$ for $i = 2$ and 3 , and $NF_1(M) = F_1(M)$; however, whether $DF_1(M) = NF_1(M)$ holds or not is still open, and this problem is called the **LBA problem**. These automata are used to study the closure properties under various operations within the families $F_i(M)$. In particular, the LBA problem is equivalent to the problem of whether $F_i(M)$ is closed under complementation.

E. Applications of Automata Theory

Finite automata are formulated as models for digital circuits with memory. Therefore there exist structural studies of finite automata; for example, methods to decompose a given finite automaton to smaller automata corresponding to a circuit and its subcircuits are studied using tsemigroup theory [14, 17].

Context-free grammars are useful in describing the syntax of programming languages. For a context-free grammar G , a **derivation tree** of $x \in L(G)$ is a tree whose terminal nodes are labeled by each symbol of x , its root is labeled by S , and other nodes are labeled by some elements of N in such a way that if for a node with $A \in N$ the nodes connected by an outgoing edge from the former have labels x_1, \dots, x_k , then (A, x_1, \dots, x_k) is a production rule of G . **Parsing** is an algorithm to construct a derivation tree of each $x \in L(G)$. Various parsing algorithms have been studied within subfamilies of context-free grammars [16] and they are used in the construction of compilers.

References

- [1] A. M. Turing, On computable numbers, with an application to the Entscheidungsproblem, Proc. London Math. Soc., (2) 42 (1936), 230–265; 43 (1937), 544–546.
- [2] E. L. Post, Finite combinatory processes—formulation 1, J. Symbolic Logic, 1 (1936), 103–105.
- [3] A. A. Markov, On the impossibility of certain algorithms in the theory of associative systems I (in Russian), Dokl. Akad. Nauk SSSR, 55 (1947), 583–586; II, Dokl. Akad. Nauk SSSR, 58 (1947), 3533356.
- [4] E. L. Post, Recursive unsolvability of a problem of Thue, J. Symbolic Logic, 12 (1947), 1–11.
- [5] S. C. Kleene, Introduction to metamathematics, Van Nostrand, 1952.
- [6] M. Davis, Computability and unsolvability, McGraw-Hill, 1958.

- [7] C. E. Shannon and J. McCarthy, Automata studies, Ann. Math. Studies, Princeton Univ. Press, 1956.
- [8] S. C. Kleene, Turing-machine computable functionals of finite types I, Logic, methodology and philosophy of science, Proc. 1960 Intern. Congr. Logic, Stanford Univ. Press, 1962, 38845: II, Proc. London Math. Soc., (3) 12 (1962), 245-258.
- [9] H. Hermes, Enumerability, decidability, computability, Springer, 1965.
- [10] N. Chomsky, Three models for the description of language, IRE Trans. on Information Theory, 2 (1956), 1133-124.
- [11] J. E. Hopcroft and J. D. Ullman, Introduction to automata theory, languages, and computation, Addison-Wesley, 1979.
- [12] S. Eilenberg, Automata, languages and machines, 2 vols., Academic Press, 1974, 1976.
- [13] J. Berstel, Transductions and context-free languages, Teubner, 1979.
- [14] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, The design and analysis of computer algorithms, Addison-Wesley, 1974.
- [15] A. V. Aho and J. D. Ullman, Principles of compiler design, Addison-Wesley, 1977.
- [16] K. Krohn and J. L. Rhodes, Algebraic theory of machines I, Prime decomposition theorem for finite semi-groups and machines, Trans. Amer. Math. Soc., 116 (1965), 450-464.
- [17] A. Nozaki, Practical decomposition of automata, Information and Control, 36 (1978), 275-291.

32 (X1.22) Automorphic Functions

A. General Remarks [1]

Let X be a (complex) \dagger analytic manifold and Γ a \dagger discontinuous group of (complex) analytic automorphisms of X . A set of \dagger holomorphic functions (without zero) on X , $\{j_\gamma(z)\}$ ($\gamma \in \Gamma$), is called a **factor of automorphy** if it satisfies the condition $j_{\gamma\gamma'}(z) = j_\gamma(\gamma'(z))j_{\gamma'}(z)$ for all $\gamma, \gamma' \in \Gamma$, $z \in X$. A \dagger meromorphic function f on X is called a (multiplicative) **automorphic function** with respect to the factor of automorphy $\{j_\gamma(z)\}$ if it satisfies the condition $f(\gamma(z)) = f(z)j_\gamma(z)$ for all $\gamma \in \Gamma$, $z \in X$. When all $j_\gamma(z)$ are identically equal to 1, i.e., when $f(z)$ is Γ -invariant, f is simply called an **automorphic function with respect to Γ** . When all j_γ are constant (and hence $\gamma \rightarrow j_\gamma$ is a "quasicharacter" of Γ), f is called a **multiplicative function**. If we denote by $J_\gamma(z)$ the functional determinant of the transformation γ , then for an integer m , an automorphic function f with respect to

Automorphic Functions

a factor of automorphy of the form $j_{\gamma}(z) = J_\gamma(z)^m$ is called an **automorphic form of weight m** . (In this case, it is customary to assume that f is holomorphic. Also, in this and most other cases, we add a suitable condition on the behavior off at the "points at infinity," when $\Gamma \backslash X$ is not compact (- 234 Kleinian Groups).)

B. The Case of One Variable [2-6]

Except for the cases where Γ is a finite group or its extension by a free Abelian group of rank ≤ 2 (\rightarrow 134 Elliptic Functions E), it is essentially enough to consider the case where X is the upper half-plane $\mathfrak{H} = \{z \in \mathbb{C} \mid \text{Im } z > 0\}$. In this case, Γ may be obtained as a \dagger discrete subgroup of the \dagger special linear group $G = SL(2, \mathbb{R})$. In the following, we restrict ourselves to the case where Γ is a \dagger Fuchsian group of the first kind (i.e., the case where the \dagger Haar measure $\mu(\Gamma \backslash \mathfrak{H})$ of the quotient space $\Gamma \backslash \mathfrak{H}$ is finite; \rightarrow 122 Discontinuous Groups). An **automorphic function** (or **Fuchsian function**) f with respect to Γ is, by definition, a meromorphic function f on \mathfrak{H} satisfying the following conditions: (i) f is Γ -invariant, i.e., $f(\gamma z) = f(z)$ for all $\gamma \in \Gamma$; (ii) f is also meromorphic around any \dagger cusp x_0 of Γ , i.e., if φ_0 is a real linear fractional transformation mapping x_0 to ∞ (e.g., $\varphi_0(z) = -1/(z - x_0)$) and if the transformation $z \rightarrow z + h$ ($h > 0$) is a generator of $\varphi_0 \Gamma_{x_0} \varphi_0^{-1}$ (where Γ_{x_0} is the \dagger stabilizer of x_0), then $f(\varphi_0^{-1}(z))$ can be expanded into a \dagger Laurent series of $q_h = \exp((2\pi i/h)z)$ (which has only finitely many terms with negative exponent) in a neighborhood $\text{Im } z > y$, of ∞ (q_h is called the **local parameter** around the cusp x_0). If we denote by \mathfrak{R}_Γ the compact \dagger Riemann surface obtained from the quotient space $\Gamma \backslash \mathfrak{H}$ by adjoining a certain (finite) number of "points at infinity" corresponding to the equivalence classes of cusps of Γ , then conditions (i) and (ii) amount to saying that f gives rise in a natural manner to a meromorphic function on \mathfrak{R}_Γ . Thus the field \mathfrak{A}_Γ of all automorphic functions with respect to Γ can be identified with the \dagger algebraic function field belonging to the Riemann surface \mathfrak{R}_Γ (\rightarrow 11 Algebraic Functions A). (While any (nonconstant) automorphic function with respect to a Fuchsian group of the first kind Γ has the real axis as its \dagger natural boundary, an automorphic function with respect to a Fuchsian group of the second kind can always be analytically extended to the lower half-plane through a neighborhood of any "ordinary point" on the real axis (\rightarrow 234 Kleinian Groups).)

Let k be an even integer. A (holomorphic) **automorphic form** (or **Fuchsian form**) f of

weight k or of dimension $-k$ with respect to Γ is, by definition, a holomorphic function f on \mathfrak{H} satisfying the following conditions: (i) For every $\sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, $f(\sigma z) = f(z) \cdot (cz + d)^k$. (In other words, $J_\sigma(z)^{-k/2} = (d\sigma(z)/dz)^{-k/2} = (cz + d)^k$ is the factor of automorphy.) (ii) f is also holomorphic around any cusp x_0 , i.e., in the above notation, we have an integral "Fourier expansion" $f(\varphi_0^{-1}(z))(d\varphi_0^{-1}(z)/dz)^{k/2} = \sum_{n=0}^{\infty} a_n q_h^n$. Note that this definition of weight is different from that given in Section A (- Sections F, G). If, moreover, the constant term a_0 in this Fourier expansion vanishes at all cusps of Γ , f is called a **cusp form**. In particular, f is a cusp form of weight 2 if and only if $f(z)dz$ gives rise in a natural manner to a \dagger differential form of the first kind on the Riemann surface \mathfrak{R}_Γ . (By a slight modification of condition (ii), we can also define an automorphic form of odd weight k . When $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in \Gamma$, condition (i) (for k odd) implies that f is identically equal to zero. So assume that $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \notin \Gamma$. Then cusps x_0 are classified into two categories, according as $\begin{pmatrix} -1 & h \\ 0 & -1 \end{pmatrix} \in \varphi_0 \Gamma_{x_0} \varphi_0^{-1}$ or not, and in the first case we should replace the power series in q_h in condition (ii) by $q_h^{1/2} \times$ (power series in q_h).) (Note that the above definition of weight is slightly different from the general one given in Section A. See also the Siegel and Hilbert modular cases (- Sections F, G).)

We denote by $\mathfrak{M}_k(\Gamma)$ (resp. $\mathfrak{S}_k(\Gamma)$) the linear space of all automorphic forms (cusp forms) of weight k with respect to Γ . Since clearly the relations $\mathfrak{M}_k \mathfrak{M}_{k'} \subset \mathfrak{M}_{k+k'}$, $\mathfrak{M}_k \mathfrak{S}_{k'} \subset \mathfrak{S}_{k+k'}$ hold, the direct sum $\mathfrak{M}(\Gamma) = \sum_k \mathfrak{M}_k(\Gamma)$ is a (commutative) \dagger graded algebra, of which $\mathfrak{S}(\Gamma) = \sum_k \mathfrak{S}_k(\Gamma)$ is an ideal. It is known that $\mathfrak{M}_k(\Gamma)$ is of finite dimension, and $d_k = \dim \mathfrak{M}_k(\Gamma)$, $d_k^0 = \dim \mathfrak{S}_k(\Gamma)$ can be determined (from the \dagger Riemann-Roch theorem) as follows:

$$d_k = 0 \quad \text{for} \quad k < 0;$$

$$d_0 = 1, \quad d_0^0 = \begin{cases} 1 & \text{for } t = 0 \\ 0 & \text{for } t > 0 \end{cases}$$

$$d_2 = \begin{cases} g & \text{for } t = 0 \\ g + t - 1 & \text{for } t > 0, \end{cases} \quad d_2^0 = g;$$

$$d_k = (k-1)(g-1) + \sum_{i=1}^s \left(\frac{k}{2} \left(1 - \frac{1}{e_i} \right) \right) + \frac{k}{2} t,$$

$$d_k^0 = d_k - t \quad \text{for} \quad k \text{ even, } \geq 4,$$

where s is the number of the equivalence classes of \dagger elliptic fixed points of Γ , e_i ($1 \leq i \leq s$) is the order of the stabilizers of these elliptic points, t is the number of the equivalence

classes of cusps of Γ , and g is the \dagger genus of the Riemann surface \mathfrak{R}_Γ . (When $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \notin \Gamma$, we can obtain a similar formula for k odd except for d_1 , d_1^0 , which, in general, cannot be determined from the Riemann-Roch theorem.)

One method of constructing automorphic forms is provided by Poincaré series as follows: Let σ_0 be a \dagger linear fractional transformation that maps the unit disk to the upper half-plane (e.g., the \dagger Cayley transformation), and put $\Gamma' = \sigma_0^{-1} \Gamma \sigma_0$; let φ be a function holomorphic on the unit disk including its boundary (e.g., a polynomial). Then for $k \geq 4$, the series

$$P_k(z) = \sum_{\sigma \in \Gamma'} \varphi(\sigma'(z)) (d\sigma'(z)/dz)^{k/2}$$

is \dagger uniformly convergent in the wide sense (i.e., uniformly convergent on every compact set) in the unit disk and expresses an automorphic cusp form of weight k with respect to Γ' , i.e., we have $P_k(\sigma_0^{-1}(z))(d\sigma_0^{-1}(z)/dz)^{k/2} \in \mathfrak{S}_k(\Gamma)$. Conversely, every element in $\mathfrak{S}_k(\Gamma)$ ($k \geq 4$) can be expressed in this form. A series of this type is called a **theta-Fuchsian series of Poincaré** (or simply **Poincaré series**).

For $k \geq 3$, we can define a (positive definite Hermitian) inner product on $\mathfrak{S}_k(\Gamma)$ as follows:

$$(f, g) = \int_F f(z) \overline{g(z)} y^{k-2} dx dy,$$

where $z = x + iy$ and F is a \dagger fundamental region of Γ in \mathfrak{H} . This is called the **Petersson metric**. Since this integral converges also for $f, g \in \mathfrak{M}_k(\Gamma)$ if one of them belongs to $\mathfrak{S}_k(\Gamma)$, we can define the orthogonal complement $\mathfrak{E}_k(\Gamma)$ of $\mathfrak{S}_k(\Gamma)$ in $\mathfrak{M}_k(\Gamma)$. As we shall see, $\mathfrak{E}_k(\Gamma)$ is generated by \dagger Eisenstein series.

Any automorphic function f with respect to Γ can be expressed as a quotient of two automorphic forms $f_1, f_2 \in \mathfrak{M}_k(\Gamma)$ for a sufficiently large k . If we put $w = f(z)$, the inverse function $z = f^{-1}(w)$ can be expressed as the quotient of two linearly independent solutions of a \dagger linear differential equation of the form $d^2 z/dw^2 = \varphi(w)z$, where $\varphi(w)$ is an \dagger algebraic function belonging to the Riemann surface \mathfrak{R}_Γ .

C. Modular Functions and Modular Forms [4, 7, 8]

The (elliptic) \dagger modular group

$$\Gamma = \Gamma(1)$$

$$= SL(2, \mathbb{Z})$$

$$= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a, b, c, d \in \mathbb{Z}, ad - bc = 1 \right\}$$

is a Fuchsian group of the first kind acting on

the upper half-plane \mathfrak{H} . The quotient space $\Gamma \backslash \mathfrak{H}$ can be compactified to get a compact Riemann surface $\mathfrak{R}_\Gamma = \Gamma \backslash \mathfrak{H} \cup \{\infty\}$ of genus zero by joining one point at infinity corresponding to the cusp ∞ , around which we take $q = e^{2\pi iz}$ as a local parameter. The dimension $d_k = \dim \mathfrak{M}_k(\Gamma)$ (k even, ≥ 2) is given as follows:

$$d_k = \begin{cases} [k/12] & \text{for } k \equiv 2 \pmod{12}, \\ [k/12] + 1 & \text{for } k \not\equiv 2 \pmod{12}, \end{cases}$$

where $[]$ is the Gauss symbol.

More generally, an automorphic function (automorphic form) with respect to the principal congruence subgroup

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}$$

is called a **modular function (modular form)** of level N . (For the number of cusps of $\Gamma(N)$, denoted by $t(N)$, and the genus of $\mathfrak{R}_{\Gamma(N)} \rightarrow 122$ Discontinuous Groups.) At any cusp of $\Gamma(N)$, the local parameter q_h in (ii) is given by $q_N = \exp(2\pi iz/N)$. For N , $k \geq 3$, the dimensions d_k , d_k^0 are given by the general formula of the preceding paragraph (including odd k). For $k \geq 3$, we define the (extended) **Eisenstein series** by

$$G_k(z; c_1, c_2, N) = \sum_{m_i \equiv c_i \pmod{N}} \frac{1}{(m_1 + m_2 z)^k},$$

where c_1, c_2 are integers such that $(c_1, c_2, N) = 1$ and the symbol \sum' denotes the summation excepting the pair $(m_1, m_2) = (0, 0)$. Then $G_k(z; c_1, c_2, N)$ depends only on $c_1, c_2 \pmod{N}$, and if we take a set $\{(c_1, c_2)\}$ such that $\{(c_1, c_2), (-c_1, -c_2)\}$ forms a complete set of representatives of the (primitive) residue classes (\pmod{N}) , then the corresponding set of Eisenstein series forms a basis of the space $\mathfrak{E}_k(\Gamma(N))$ ($k \geq 3$), whence we obtain $\dim \mathfrak{E}_k = d_k - d_k^0 = t(N)$ ($k \geq 3$) (for details, including the case $k=1, 2$, see E. Hecke, *Abh. Math. Sem. Univ. Hamburg*, 5 (1927)).

The Fourier coefficient a , of the Eisenstein series can be calculated easily, and we get an estimate $a_k = O(n^{k-1+\varepsilon})$ for $k \geq 2$, where ε is an arbitrary positive number. On the other hand, for cusp forms, Hecke (*Abh. Math. Sem. Univ. Hamburg*, 5 (1927)) gave $a_n = O(n^{k/2})$. In an attempt to improve this estimate, H. D. Kloosterman was led to consider the sum

$$K(u, v, q) = \sum_{\substack{x \pmod{q} \\ (x, q) = 1}} \exp\left(\frac{2\pi i}{q} \left(ux + \frac{v}{x}\right)\right) \quad (u, v, q \in \mathbb{Z}),$$

called the **Kloosterman sum**, which is also related to the arithmetic of quadratic forms. Using A. Weil's estimate $|K(u, v, p)| \leq 2\sqrt{p}$

(where p is an odd prime, $(u, p) = (v, p) = 1$), based on the analog of the Riemann hypothesis, we obtain $a_k = O(n^{k/2-1/4+\varepsilon})$. The estimate $a_n = O(n^{k/2-1/2+\varepsilon})$ is a weak form of what is known as the Ramanujan-Petersson conjecture (\rightarrow Section D). In the case $N=1$, we obtain the classical Eisenstein series $G_k(z) = G_k(z; 0, 0, 1)$ (k even, ≥ 4), from which we define the modular forms $g_2(z) = 60G_4(z)$, $g_3(z) = 140G_6(z)$, and $A(z) = g_2^3 - 27g_3^2$ of weight 4, 6, and 12, respectively. If we denote by $\wp = \wp(u; 1, z)$ the Weierstrass \wp -function with the fundamental period $(1, z)$ we have the relation $\wp'^2 = 4\wp^3 - g_2\wp - g_3$, and $A(z)$ is the \dagger discriminant of the cubic polynomial appearing in this relation. It is known that every modular form can be expressed uniquely as a polynomial in g_2 and g_3 , or in other words, we have $\mathfrak{M}(\Gamma(1)) \cong \mathbb{C}[g_2, g_3]$. The polynomial in g_2, g_3 expressing a modular form is, moreover, **isobaric**, i.e., consists of terms of the form $cg_2^\mu g_3^\nu$, $c \in \mathbb{C}$, where $2\mu + 3\nu$ is a constant called the **weight** of the isobaric polynomial. Also, $A(z)$ is a cusp form of the smallest weight, and the ideal of all cusp forms, $\mathfrak{S}(\Gamma(1))$, is a principal ideal in $\mathfrak{M}(\Gamma(1))$ generated by $A(z)$. The Fourier expansion of the Eisenstein series $G_k(z)$ is given as follows:

$$G_k(z) = (2\pi)^k \frac{1}{k!} \left(B_k + (-1)^{k/2} 2k \sum_{v=1}^{\infty} v^{k-1} \frac{q^v}{1-q^v} \right),$$

where $q = e^{2\pi iz}$ and B_k is the \dagger Bernoulli number. The discriminant $A(z)$ is expressed as an infinite product as follows:

$$\Delta(z) = (2\pi)^{12} q \prod_{v=1}^{\infty} (1 - q^v)^{24}.$$

If we put $J(z) = g_2^3/\Delta$, the function $J(z)$ is a modular function and gives an \dagger analytic isomorphism of the Riemann surface $\mathfrak{R}_{\Gamma(1)} = \Gamma(1) \backslash \mathfrak{H} \cup \{\infty\}$ onto the Riemann sphere $\mathbb{C} \cup \{\infty\}$ (which maps ζ_3, ζ_4, ∞ to 0, 1, ∞ , respectively). Hence the field of modular functions $\mathfrak{R}_{\Gamma(1)}$ is a rational function field $\mathbb{C}(J)$ generated by the function J . The analytic isomorphism class of \dagger complex tori of dimension 1 (= \dagger elliptic curves) $E_{(\omega_1, \omega_2)} = \mathbb{C}/(\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2)$ with the fundamental period (ω_1, ω_2) is uniquely determined by the $\Gamma(1)$ -equivalence class of the **modulus** $\tau = \omega_2/\omega_1$ ($\in \mathfrak{H}$), and hence by the value $J(\tau)$ of the function J . This is the historical origin of the name of modular function [4].

As an example of a modular function of level 2, we have the λ -function:

$$\lambda(z) = \frac{\wp((1+z)/2) - \wp(z/2)}{\wp(1/2) - \wp(z/2)},$$

which gives an analytic isomorphism of the Riemann surface $\mathfrak{R}_{\Gamma(2)} = \Gamma(2) \backslash \mathfrak{H} \cup$ (three

points) to the Riemann sphere $\mathbb{C}U \setminus \{\infty\}$ (mapping 0, 1, ∞ to 1, ∞ , 0, respectively; this property is used in a proof of the \dagger Picard theorem). Hence we again have $\mathfrak{R}_{\Gamma(2)} = \mathbb{C}(\lambda)$. The following relation holds between J and λ :

$$J(z) = \frac{4}{2} \frac{(1 - \lambda + \lambda^2)^3}{7 \lambda^2 (1 - \lambda)^2}$$

In general, the field $\mathfrak{R}_{\Gamma(N)}$ of modular functions of level N is generated by $(g_2/g_3) \cdot \wp((a_1 + a_2 z)/N; 1, z)$ ($a_1, a_2 \in \mathbb{Z}$).

D. The Hecke Ring and Its Representation [S-1 13]

In general, let $\tilde{\Gamma}$ be a group, Γ a subgroup of $\tilde{\Gamma}$, and suppose that, for every $\sigma \in \tilde{\Gamma}$, $\sigma\Gamma\sigma^{-1}$ and Γ are \dagger commensurable, which amounts to saying that every double coset $\Gamma\sigma\Gamma$ is a union of a finite number of left or right cosets of Γ . Then in the free module generated by all double cosets $\Gamma\sigma\Gamma$ ($\sigma \in \tilde{\Gamma}$), we can define a bilinear associative product as follows:

$$\Gamma\sigma\Gamma \cdot \Gamma\tau\Gamma = \sum_{\rho \in \Gamma} m(\sigma, \tau; \rho) \Gamma\rho\Gamma,$$

where, writing $\Gamma\sigma\Gamma = \bigcup_i \Gamma\sigma_i$, $\Gamma\tau\Gamma = \bigcup_j \Gamma\tau_j$, we denote by $m(\sigma, \tau; \rho)$ the number of pairs (i, j) such that $\Gamma\sigma_i\tau_j = \Gamma\rho$. The associative ring thus obtained is called the **Hecke ring** (or **Hecke algebra**) and is denoted by $\mathcal{H}(\tilde{\Gamma}, \Gamma)$. (If $\tilde{\Gamma}$ is a \dagger topological group and Γ is an open compact subgroup, the Hecke ring $\mathcal{H}(\tilde{\Gamma}, \Gamma)$ can also be interpreted as the ring of all \mathbb{Z} -valued continuous functions on $\tilde{\Gamma}$ with compact support provided with a \dagger convolution product; \rightarrow 29 Associative Algebras C.)

As an example, setting $\tilde{\Gamma} = GL^+(2, \mathbb{Q})$, $\Gamma = SL(2, \mathbb{Z})$, we obtain a Hecke ring $\mathcal{H}(\tilde{\Gamma}, \Gamma)$. In virtue of the theory of \dagger elementary divisors, a complete set of representatives of $\Gamma \backslash \tilde{\Gamma} / \Gamma$ is given by

$$\left\{ \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \mid a_1, a_2 \in \mathbb{Q}, a_1, a_2 > 0, a_1 \mid a_2 \right\}.$$

We write $T(a_1, a_2) = \Gamma \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \Gamma$. From the relation $T(a_1 a_2, a_1 a_2) T(a_1^{-1}, a_2^{-1}) = T(a_1, a_2)$ follows the commutativity of $\mathcal{H}(\tilde{\Gamma}, \Gamma)$. Furthermore, if we put, for a positive integer n ,

$$T_n = T(n) = \sum_{\substack{a_1, a_2 \in \mathbb{Z} \\ a_1 a_2 = n}} T(a_1, a_2),$$

then the following multiplication formula holds:

$$T(n) \cdot T(m) = \sum_{d \mid n, m} d T(d, d) \cdot T(nm/d^2).$$

We obtain a representation of $\mathcal{H}(\tilde{\Gamma}, \Gamma)$ in $\mathfrak{M}_k(\Gamma)$ as follows. For $T = \Gamma\sigma\Gamma = \bigcup \Gamma\sigma_i\Gamma$

$\mathcal{H}(\tilde{\Gamma}, \Gamma)$ and $f \in \mathfrak{M}_k(\Gamma)$ we put

$$(f|T)(z) = \sum_i f(\sigma_i(z)) (c_i z + d_i)^{-k} \det(\sigma_i)^l,$$

where $\sigma_i = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$ and l is a fixed integer. This representation leaves $\mathfrak{S}_k(\Gamma)$ and $\mathfrak{E}_k(\Gamma)$ invariant. In particular, the representation of T_n (with $l = k - 1$), called the **Hecke operator** (Hecke, *Math. Ann.*, 114 (1937)), is \dagger Hermitian with respect to the Petersson metric.

Following Hecke, we associate with every modular form $f(z) = \sum_{n=0}^{\infty} a_n q^n$ ($q = e^{2\pi i z}$) a \dagger Dirichlet series $\varphi(s) = \sum_{n=1}^{\infty} a_n n^{-s}$. Since $a_n = O(n^{k-1+\epsilon})$ for $f \in \mathfrak{M}_k$ ($k \geq 2$), $\varphi(s)$ is absolutely convergent in the half-plane $\text{Re } s > k$. The conditions for f to be a modular form of weight k are equivalent to the following conditions for φ : (s-k) $\varphi(s)$ can be extended to an entire function of finite genus (actually, of genus 1) and, if we put $R(s) = (2\pi)^{-s} \Gamma(s) \varphi(s)$, $R(s)$ satisfies a functional equation of the form $R(k-s) = (-1)^{k/2} R(s)$. A correspondence between f and φ that satisfies these conditions is one-to-one; in fact we have $a_n = (-1)^{k/2} \text{Res}_{s=k} R(s)$, and the function $g(x) = f(ix) - a_0(x > 0)$ and $R(s)$ are related by the \dagger Mellin transform as follows:

$$R(s) = \int_0^{\infty} g(x) x^{s-1} dx,$$

$$g(x) = \frac{1}{2\pi i} \int_{\text{Res}=\sigma_0} R(s) x^{-s} ds.$$

This correspondence between Dirichlet series and automorphic forms can be generalized to the case of a certain discrete subgroup of $SL(2, \mathbb{R})$, which is not necessarily of the first kind (Hecke, *Math. Ann.*, 112 (1936); [10]). It is also known that if the functional equations for $\sum_{n=1}^{\infty} \chi(n) a_n n^{-s}$ hold for sufficiently many characters χ , then $f = \sum_{n=1}^{\infty} a_n q^n$ is a cusp form with a character for some congruence subgroup (Weil, *Math. Ann.*, 168 (1967); [12]).

Now, suppose that a linear subspace \mathfrak{M} of $\mathfrak{M}_k(\Gamma)$ is invariant under all T_n ($n = 1, 2, \dots$), and let (f_1, \dots, f_κ) ($\kappa = \dim \mathfrak{M}$) be a basis of \mathfrak{M} . If we denote by $\lambda(n)$ the $\kappa \times \kappa$ matrix representing T_n in this basis, it can be shown that there exist $\kappa \times \kappa$ (complex) matrices $B^{(i)}$ ($1 \leq i \leq \kappa$) and $\lambda(0)$ such that we have

$$F(z) = \sum_{n=0}^{\infty} q^n \lambda(n) = \sum_{i=1}^{\kappa} f_i(z) B^{(i)}.$$

Similarly we associate with $F(z)$ a matrix-valued Dirichlet series

$$\Phi(s) = \sum_{n=1}^{\infty} n^{-s} \lambda(n).$$

Then from the multiplicative property of the $\lambda(n)$, we obtain the following \dagger Euler product

expression of $\Phi(s)$:

$$\Phi(s) = \prod_p (I_\kappa - \lambda(p)p^{-s} + p^{\kappa-1-2s}I_\kappa)^{-1},$$

where I_κ denotes the unit matrix of degree κ . In particular, when $\dim \mathfrak{M} = 1$, i.e., when $\mathfrak{M} = C f$, where f is an eigenfunction of all Hecke operators T_n , the corresponding Dirichlet series $\varphi(s)$ has an Euler product of the following form:

$$\varphi(s) = \prod_p (1 - \lambda_p p^{-s} + p^{\kappa-1-2s})^{-1}.$$

For instance, the Dirichlet series corresponding to the Eisenstein series G_k is $(2(2\pi i)^k (k-1)!) \zeta(s) \zeta(s-k+1)$ (where ζ is the Riemann zeta function), which has a well-known Euler product. Since $\mathfrak{S}_{12} = CA$, the Dirichlet series corresponding to $\Delta(z)$ also has an Euler product. (This is part of the Ramanujan conjecture. Ramanujan also conjectured that the quadratic polynomial $1 - \lambda_p X + p^{11} X^2$ appearing in this Euler product has imaginary roots, i.e., $|\lambda_p| < 2p^{1/2}$.) Generalizing the Ramanujan conjecture, Petersson conjectured that all eigenvalues of the Hecke operator T_p in \mathfrak{S}_k have absolute value $\leq 2p^{(k-1)/2}$ (Ramanujan-Petersson conjecture). P. Deligne proved this conjecture for $k \geq 2$, and his method extends to the case when Γ is a congruence subgroup (*Publ. Math. Inst. HES*, 53 (1974)). For $k=1$, see Deligne and J.-P. Serre (*Ann. Sci. Ecole Norm. Sup.*, 7 (1974)). It may be noted that this conjecture was proved earlier for $k=2$ (and for almost all p) by M. Eichler (1954) and by G. Shimura (1958) (\rightarrow 450 Zeta Functions M).

E. The Case of Many Variables [1, 15, 161]

Let X be a bounded domain in C^N and Γ a discontinuous group of analytic automorphisms of X . The Poincaré series of weight $(\rightarrow$ Section A) $m \geq 2$, defined similarly to the case of the unit disk, converges normally in X and expresses an automorphic form of weight m . If $\Gamma \backslash X$ is compact, let (f_1, f_κ) be a basis of the space of Poincaré series of weight m . Then for a sufficiently large m (which is a multiple of the order of Γ_x for all $x \in X$), the map $X \ni z \rightarrow (f_1(z), f_\kappa(z))$ defines in a natural manner a projective embedding of the quotient space $\Gamma \backslash X$ into $P^{\kappa-1}(C)$, which actually gives an analytic isomorphism of $\Gamma \backslash X$ onto a normal projective algebraic variety. It follows that the field of automorphic functions with respect to Γ is an algebraic function field of dimension N , and that every automorphic function can be written as the quotient of two Poincaré series of the same weight. In the case where $\Gamma \backslash X$ is not compact, we first discuss some examples.

F. Siegel Modular Functions [15, 17, 18]

The Siegel upper half-space \mathfrak{H}_n of degree n (or Siegel space of degree n) is, by definition, the space of all $n \times n$ complex symmetric matrices $Z = X + iY$ with the imaginary part $Y > 0$ (positive definite). (\mathfrak{H}_n is analytically equivalent to a symmetric bounded domain.) The group of all (complex) analytic automorphisms of \mathfrak{H}_n is given by the real (projective) symplectic group $Sp(n, \mathbf{R}) / \{\pm I_{2n}\}$, which acts transitively on \mathfrak{H}_n by $Z \rightarrow \sigma(Z) = (AZ + B)/(CZ + D)$ for $\sigma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n, \mathbf{R})$. The subgroup $\Gamma_n = Sp(n, \mathbf{Z})$ consisting of integral matrices, or the corresponding group of linear fractional transformations, is called the Siegel modular group of degree n . Γ_n is a discontinuous group of the first kind acting on \mathfrak{H}_n .

A Siegel modular form f of weight k or of dimension $-k$ is, by definition, a holomorphic function on \mathfrak{H}_n satisfying the following conditions: (i) for $\sigma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_n$, $f(\sigma(Z)) = f(Z) \det(CZ + D)^k$; (ii) f has an integral Fourier expansion of the form $f(Z) = \sum_{T \geq 0} a_T e^{2\pi i \text{tr}(TZ)}$, where T runs over all half-integral positive semidefinite symmetric matrices of degree n . (Note that the Jacobian determinant $J_\sigma(Z)$ of $\sigma = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is given by $J_\sigma(Z) = \det(CZ + D)^{-k-1}$.) For $n \geq 2$, condition (ii) is superfluous (M. Koecher, *Math. Z.*, 59 (1954)). We denote by $\mathfrak{M}_k^{(n)} = \mathfrak{M}_k(\Gamma_n)$ the space of all Siegel modular forms of degree n and weight k . When we write $\mathfrak{H}_n \ni Z = \begin{pmatrix} Z_1 & 3 \\ 3 & z \end{pmatrix}$ with $Z_1 \in \mathfrak{H}_{n-1}$, $3 \in C^{n-1}$, $z \in \mathfrak{H}_1$, and

$$\mathfrak{M}_k^{(n)} \ni f(Z) = \sum_{n=0}^{\infty} a_n(Z_1, 3) e^{2\pi i n z},$$

it turns out that $a_0(Z_1, 3)$ depends only on Z_1 and, writing it as $f_1(Z_1)$, we have $f_1 \in \mathfrak{M}_k^{(n-1)}$. The mapping $\Phi: f \rightarrow f_1$ thus defined is a linear mapping from $\mathfrak{M}_k^{(n)}$ into $\mathfrak{M}_k^{(n-1)}$, which is surjective if k is even and $> 2n$ (H. Maass, *Math. Ann.*, 123 (1951)). We denote the kernel of Φ by $\mathfrak{S}_k^{(n)}$ and call $f \in \mathfrak{S}_k^{(n)}$ a cusp form (viewing \mathfrak{H}_{n-1} as a cusp of \mathfrak{H}_n). For $f \in \mathfrak{M}_k^{(n)}$, the following three conditions are equivalent: (a) $f \in \mathfrak{S}_k^{(n)}$; (b) $f(Z) = \sum_{T > 0} a_T e^{2\pi i \text{tr}(TZ)}$; (c) $|f(Z) \det(Y)^{k/2}|$ is bounded. We have $\mathfrak{M}_k^{(n)} = \{0\}$ if $k < 0$ or if both k and n are odd, and $\mathfrak{M}_0^{(n)} = C$. In general, $\dim \mathfrak{M}_k^{(n)}$ is finite and $= O(k^{n(n+1)/2})$ ($k \rightarrow \infty$). The graded algebra $\mathfrak{M}(\Gamma_n) = \sum_{k=0}^{\infty} \mathfrak{M}_k^{(n)}$ is finitely generated, and its Krull dimension is $n(n+1)/2 + 1$. In particular, for $n=2$ the structure of the graded algebra $\mathfrak{M}(\Gamma_2)$ is determined explicitly. (The even part of this algebra is isomorphic to the

polynomial ring of four variables.) Also, $\dim \mathfrak{M}_k^{(2)}$ and the singularities of the quotient variety $\Gamma_2 \backslash \mathfrak{H}_2$ are known explicitly (J. I. Igusa, *Amer. J. Math.*, 84 (1962), 86 (1964); U. Christian).

As an example of Siegel modular forms, we have the **Eisenstein-Poincaré series**, defined as follows:

$$E_{S,k}(Z) = \sum_{\sigma \in \Gamma_n(S) \backslash \Gamma_n} e^{2\pi i \operatorname{tr}(S\sigma(z))} \det(CZ + D)^{-k},$$

where S is an $n \times n$ rational symmetric matrix ≥ 0 , and we put

$$\Gamma_n(S) = \left\{ \begin{pmatrix} U & T^t U^{-1} \\ 0 & {}^t U^{-1} \end{pmatrix} \in \Gamma_n \mid {}^t U S U = S, \right. \\ \left. e^{2\pi i \operatorname{tr}(ST)} = \det(U)^k \right\}$$

This series is convergent for $k > n + \operatorname{rank} S + 1$ if $\operatorname{rank} S < n$ and for $k > 2n$ if $S > 0$, and the totality of $E_{S,k}$ (those series with $S > 0$) spans $\mathfrak{M}_k^{(n)}(\mathfrak{S}_k^{(n)})$ (Maass, *Math. Ann.*, 123 (1951)). †Theta series defined by integral quadratic forms are examples of Siegel modular forms (with a certain level) and are of great significance in the arithmetic of quadratic forms (C. L. Siegel, *Ann. Math.*, (2) 36 (1935)).

The quotient space $\Gamma_n \backslash \mathfrak{H}_n$ can be compactified as follows (I. Satake and W. Baily [15]): There exists a positive integer k_0 such that, for any multiple k of k_0 , any basis of $\mathfrak{M}_k^{(n)}$ defines in a natural manner a one-to-one biholomorphic projective embedding of $\Gamma_n \backslash \mathfrak{H}_n$, of which the image is a Zariski open set of a normal projective algebraic variety. Since the structure of this projective variety is independent of k , we denote it simply by $\Gamma_n \backslash \mathfrak{H}_n$. In other words, $\Gamma_n \backslash \mathfrak{H}_n$ is the projective variety $\operatorname{Proj} \mathfrak{M}(\Gamma_n)$ associated with the graded algebra $\mathfrak{M}(\Gamma_n)$. Then we have $\Gamma_n \backslash \mathfrak{H}_n = \bigcup_{r=0}^n \Gamma_r \backslash \mathfrak{H}_r$. When $n \geq 2$ for $r < n$, $\Gamma_r \backslash \mathfrak{H}_r$ is of codimension ≥ 2 in $\Gamma_n \backslash \mathfrak{H}_n$, so that (by virtue of †Hartogs's continuation theorem) the conditions at the points at infinity in the definitions of modular functions and forms become superfluous. In fact, for $n \geq 2$, if we define a **Siegel modular function of degree n** simply as a Γ_n -invariant meromorphic function on \mathfrak{H}_n , then it can automatically be extended to a meromorphic function on the compactification $\Gamma_n \backslash \mathfrak{H}_n$, and hence be expressed as the quotient of two modular forms of the same weight. It also follows that the field of all Siegel modular functions of degree n is an algebraic function field of dimension $n(n+1)/2$. (These results can also be obtained from the pseudoconcavity of $\Gamma_n \backslash \mathfrak{H}_n$ without using compactification (A. Andreotti and H. Grauert, *Nachr. Akad. Wiss. Göttingen*.) For $n = 2$, the field of Siegel modular

functions is a rational function field. On the other hand, for $\mathbb{F}_1 \equiv 1 \pmod{8}$, $n > 9$, the field of Siegel modular functions is not rational (E. Freitag, *Abh. Math. Sem. Univ. Hamburg* (1978)).

Let Γ be a subgroup of Γ_n of finite index. Then $\Gamma \backslash \mathfrak{H}_n$ can be compactified in the same way as above [23]. However the singularities of $\Gamma \backslash \mathfrak{H}_n$ are extraordinarily complicated (Igusa [24]). When Γ is the principal congruence subgroup $\Gamma_n(N)$ of level $N > 2$, the blowing-up of $\Gamma_n(N) \backslash \mathfrak{H}_n$ along the cusps is nonsingular for $n = 2, 3$ (Igusa, *Math. Ann.*, 168 (1967)). Following Igusa's idea, it is possible to construct an explicit nonsingular compactification of $\Gamma \backslash D$, in which D is a bounded symmetric domain and Γ is a suitable arithmetic subgroup of $\operatorname{Aut}(D)$ ([19]; see also Satake, *Bull. Amer. Math. Soc.*, 79 (1973)). For $D = \mathfrak{H} \times \mathfrak{H}$ and Γ the Hilbert modular group (\rightarrow Section G), F. Hirzebruch has given explicit nonsingular compactifications of $\Gamma \backslash D$ (*Enseignement Math.*, 19 (1973)).

The Hecke theory (\rightarrow Section D) can also be extended to the Siegel modular case. For any integer $m > 0$, let $S_m = S_m^{(n)}$ be the set of $2n \times 2n$ integral matrices M such that ${}^t M J M = m J$,

where $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$ and I_n is the identity

matrix of degree n . We define the Hecke operator $T(m)$ on $\mathfrak{M}_k^{(n)}$ by

$$T(m)f = m^{nk-n(n+1)/2} \sum_{\begin{pmatrix} A & B \\ C & D \end{pmatrix} = M \in V(m)} \det(CZ + D)^{-k} f(MZ),$$

where $f \in \mathfrak{M}_k^{(n)}$, $S_m = \bigcup_{M \in V(m)} \Gamma_n M$, and $V(m)$ is a complete system of representatives of $\Gamma_n \backslash S_m$. Then for any prime number p , the formal operator series $D_p(X) = \sum_{n=0}^{\infty} T(p^n) X^n$ is formally equal to a rational function in X :

$$D_p(X) = \frac{P_p(X)}{Q_p(X)},$$

in which $P_p(X)$ and $Q_p(X)$ are polynomials in X of degree $2^{n-2} - 2$ and 2^n , respectively, with coefficients in the ring of Hecke operators (Satake, *Publ. Math. Inst. HES*, 18 (1964)). For $n = 2$ we have (Shimura, *Proc. Nat. Acad. Sci. US*, 49 (1963))

$$D_p(X) = (1 - p^{2k-4} X^2) \{ 1 - T(p)X + (T(p)^2 - T(p^2) - p^{2k-4} X^2 - p^{2k-3} T(p)X^3 + p^{4k-6} X^4) \}^{-1}.$$

Let $f(Z) \in \mathfrak{M}_k^{(2)}$ be a Siegel modular form of degree 2 such that $T(m)f = \lambda(m)f$ for all $m \geq 1$. We define $L_f(s)$ by

$$L_f(s) = \zeta(2s - 2k + 4) \sum_{m=1}^{\infty} \frac{\lambda(m)}{m^s},$$

in which $\zeta(s)$ is the Riemann zeta function. It is known that $L_f(s)$ can be meromorphically continued to the entire plane and satisfies the functional equation

$$\Psi_f(2k-2-s) = \Psi_f(s),$$

where $Y(s) = (2\pi)^{-2s} \Gamma(s) \Gamma(s-k+2) L_f(s)$ and the function $(s-k)(s-k+2) \Psi_f(s)$ is entire (A. Andrianov, *Trudy Math. Inst. Steklov.*, 112 (1971)).

G. Hilbert Modular Functions

Following a suggestion of Hilbert, O. Blumenthal studied a generalization of modular functions of the following type: Let K be a totally real algebraic number field of finite degree, and let $K^{(1)}, \dots, K^{(n)}$ be the conjugates of K , where $n = [K : \mathbb{Q}]$. ("Totally real" means that all the $K^{(i)}$ are real.) For $\alpha \in K$, we denote by $\alpha^{(i)} \in K^{(i)}$ the i th conjugate of α . Let \mathfrak{D} be the +Principal order of K , and consider the group

$$\Gamma_{\mathfrak{D}} = SL(2, \mathfrak{D}) \\ = \left\{ \sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \mid \alpha, \beta, \gamma, \delta \in \mathfrak{D}, \alpha\delta - \beta\gamma = 1 \right\}.$$

We define the action of $\Gamma_{\mathfrak{D}}$ on the n -fold product of the upper half-plane $\mathfrak{H}^n = \{z = (z_1, \dots, z_n) \mid z_i \in \mathfrak{H}\}$ by

$$a(z) = \left(\frac{\alpha^{(1)} z_1 + \beta^{(1)}}{\gamma^{(1)} z_1 + \delta^{(1)}}, \dots, \frac{\alpha^{(n)} z_n + \beta^{(n)}}{\gamma^{(n)} z_n + \delta^{(n)}} \right)$$

for $\sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$; then $\Gamma_{\mathfrak{D}}$ becomes a discontinuous group of the first kind ($\Gamma_{\mathfrak{D}}$ can also be considered as an irreducible discrete subgroup of $SL(2, \mathbb{R})^n$). The group $\Gamma_{\mathfrak{D}}$ is called the **Hilbert modular group** of K (in the strict sense). (The definition is sometimes modified by replacing the condition $\alpha\delta - \beta\gamma = 1$ by, say, $\alpha\delta - \beta\gamma = (\text{totally positive unit})$.) If the class number of K is h , the quotient space $\Gamma_{\mathfrak{D}} \backslash \mathfrak{H}^n$ can be compactified by adjoining h points at infinity. Therefore, if $n \geq 2$, a **Hilbert modular function** can be defined as a meromorphic function on \mathfrak{H}^n invariant under $\Gamma_{\mathfrak{D}}$; and similarly a **Hilbert modular form** f of weight k or of dimension $-k$ as a holomorphic function f on \mathfrak{H}^n such that $(f|_{\sigma})(z) = f(\sigma(z)) \prod_{i=1}^n (\gamma^{(i)} z_i + \delta^{(i)})^{-k} = f(z)$ for all $\sigma \in \Gamma_{\mathfrak{D}}$. (In the latter case, f is holomorphic at all cusps, i.e., for every $\sigma \in SL(2, K)$, $f|_{\sigma}$ has an integral Fourier expansion.) For Hilbert modular functions and forms, results quite similar to those for the case $n = 1$ or the case of Siegel modular groups have been obtained (Kloosterman, Maass, K. B. Gundlach, H. Klingen).

H. Further Generalizations

As further generalizations of the notion of modular function, we have Hilbert-Siegel modular functions (I. I. Pyatetskii-Shapiro, Baily, Christian [20], where we can find some 150 references), Hermitian modular functions (H. Braun, Klingen), etc. For the most general case, i.e., for an arithmetically defined discontinuous group acting on a symmetric bounded domain, a unified theory of automorphic functions has been established in the works of Pyatetskii-Shapiro [21, 22] and Baily and A. Borel [23, 24].

On the other hand, exactly as in the classical theory, where the elliptic modular function gave an invariant of 1-dimensional complex tori, generalized modular functions can be viewed as giving an analytic invariant of a certain family of (polarized) Abelian varieties. From this point of view, a deep number-theoretic (and algebrogeometric) study of automorphic functions (initiated by Hecke and Eichler) has been substantially carried forward by the work of Shimura (see the series of his papers in *Ann. Math.* starting from vol. 70 (1959); see also [9, 24, 25]).

From the analytic point of view, the theory of automorphic functions is closely connected with the unitary representation of G in the space $L_2(\Gamma \backslash G)$ (or its adelic analog) [26, 27] (for adelic analogs \rightarrow [12, 131]). In this respect, the trace formula of A. Selberg [28], generalizing the +Poisson summation formula, is of fundamental importance; and actually it can be used effectively for calculations of the dimension of the space of automorphic forms and of the trace of Hecke operators (R. P. Langlands, *Amer. J. Math.*, 85 (1963); H. Shimizu, *Ann. Math.*, (2) 77 (1963), *J. Fac. Sci. Univ. Tokyo*, 10 (1963); also [24]). When $X = G/K$ is a symmetric domain, we can define, for any representation ρ of K , a (matrix-valued) canonical automorphy factor, by which we define (vector-valued) automorphic forms with respect to a discrete subgroup Γ of G , and under a further condition (say, Γ free, $\Gamma \backslash X$ compact, and the highest weight of ρ sufficiently large) we obtain a formula for the dimension of the space of such automorphic forms in terms of the arithmetic genus of $\Gamma \backslash X$ and certain numbers related to the "dual" $X_u = G_u/K$ and the representation ρ [15, 29, 30].

References

- [1] Sém. H. Cartan 6, Théorie des fonctions de plusieurs variables; Fonctions automorphes et espaces analytiques, Ecole Norm. Sup., 1953–1954.

- [2] M. Eichler, Einführung in die Theorie der algebraischen Zahlen und Funktionen, Birkhäuser, 1963; English translation, Introduction to the theory of algebraic numbers and functions, Academic Press, 1966.
- [3] R. Fricke and F. Klein, Vorlesungen über die Theorie der automorphen Funktionen, Teubner, I, 1897; II, 1901; second edition, 1926.
- [4] F. Klein and R. Fricke, Vorlesungen über die Theorie der elliptischen Modulfunktionen, Teubner, I, 1890; II, 1892.
- [5] J. Lehner, Discontinuous groups and automorphic functions, Amer. Math. Soc. Math. Surveys, 1964.
- [6] H. Poincaré, Mémoire sur les fonctions fuchsiennes, Acta Math., 1 (1882), 193-294 (Oeuvres, Gauthier-Villars, 1916, vol. 2, 167-257).
- [7] A. P. Ogg, Modular forms and Dirichlet series, Benjamin, 1969.
- [8] G. Shimura, Introduction to the arithmetic theory of automorphic functions, Publ. Math. Soc. Japan, Iwanami Shoten and Princeton Univ. Press, 1971.
- [9] M. Eichler, Quadratische Formen und Modulfunktionen, Acta Arith., 4 (1958), 217-239.
- [10] E. Hecke, Mathematische Werke, Vandenhoeck & Ruprecht, 1959.
- [11] E. Hecke, Dirichlet series, modular forms and quadratic forms, Lecture notes, Institute for Advanced Study, Princeton, 1938.
- [12] H. Jacquet and R. P. Langlands, Automorphic forms on $GL(2)$, Lecture notes in math. 114, Springer, 1970.
- [13] Automorphic forms, representations and L -functions, Amer. Math. Soc. Proc. Symposia in Pure Math., 33 (1979).
- [14] Modular functions of one variable I-VI, Lecture notes in math., Springer, 320 (1973); 349 (1973); 350 (1973); 476 (1975); 601 (1977); 627 (1977).
- [15] Sém. H. Cartan 10, Fonctions automorphes, Ecole Norm. Sup., 1957-1958.
- [16] C. L. Siegel, Analytic functions of several complex variables, Lecture notes, Institute for Advanced Study, Princeton, 1950.
- [17] H. Maass, Siegel's modular forms and Dirichlet series, Lecture notes in math. 216, Springer, 1971.
- [18] C. L. Siegel, Einführung in die Theorie der Modulfunktionen n-ten Grades, Math. Ann., 116 (1939), 617-657 (Gesammelte Abhandlungen, Springer, 1966, vol. 2, 97-137).
- [19] D. Mumford, et al., Smooth compactification of locally symmetric varieties, Math. Sci. Press, 1975.
- [20] U. Christian, Zur Theorie der Hilbert-Siegelschen Modulfunktionen, Math. Ann., 152 (1963), 275-341.
- [21] I. I. Pyatetskii-Shapiro (Pjateckii-Šapiro), Automorphic functions and the geometry of classical domains, Gordon & Breach, 1969. (Original in Russian, 1961.)
- [22] I. I. Pyatetskii-Shapiro (Pjateckii-Šapiro), Arithmetic groups in complex domains, Russian Math. Surveys, 19 (1964), 83-109. (Original in Russian, 1964.)
- [23] W. L. Baily and A. Borel, Compactification of arithmetic quotients of bounded symmetric domains, Ann. Math., (2) 84 (1966), 442-528.
- [24] Algebraic groups and discontinuous subgroups, Amer. Math. Soc. Proc. Symposia in Pure Math., 9 (1966).
- [25] G. Shimura, Automorphic functions and number theory, Lecture notes in math. 54, Springer, 1968.
- [26] I. M. Gel'fand, M. I. Graev, and I. I. Pyatetskii-Shapiro (Pjateckii-Šapiro), Generalized functions 6, Representation theory and automorphic functions, Saunders, 1969. (Original in Russian, 1966.)
- [27] Harish-Chandra, Automorphic forms on semisimple Lie groups, Lecture notes in math. 62, Springer, 1968.
- [28] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series, J. Indian Math. Soc., 20 (1956), 47-87.
- [29] Y. Matsushima and S. Murakami, On vector bundle valued harmonic forms and automorphic forms on symmetric Riemannian manifolds, Ann. Math., (2) 78 (1963), 365-416.
- [30] Y. Matsushima and S. Murakami, On certain cohomology groups attached to Hermitian symmetric spaces, Osaka J. Math., 2 (1965), 1-35.
- [31] R. Tsushima, A formula for the dimensions of spaces of Siegel cusp forms of degree three, Amer. J. Math., 102 (1980), 933-977.

33 (1.5) Axiomatic Set Theory

A. General Remarks

Axiomatic set theory pursues the goal of reestablishing the essentials of G. Cantor's rather intuitive \dagger set theory by axiomatic constructions consistent with modern theories of the foundations of mathematics.

A system of axioms for set theory was first given by E. Zermelo [47], and was completed by A. Fraenkel [8]. J. von Neumann [30] expressed it in \dagger symbolic logic, gave a formal generalization, and eliminated ambiguous concepts. P. Bernays and K. Gödel [2, 11]

further refined and simplified von Neumann's formulation. The theories based on the systems before and after the formal generalization are called **Zermelo-Fraenkel set theory** (ZF) and **Bernays-Gödel set theory** (BG), respectively.

B. Zermelo-Fraenkel Set Theory

ZF is a formal system expressed in first-order predicate logic with the predicate symbol = (equality) and based on the following axioms 1-9 (- 411 Symbolic Logic). These axioms do not contain any predicate symbol other than \in , where $x \in y$ is read "x is an element of y." Any formula containing only \in as a predicate symbol is called a **set-theoretic formula**. Following the usual convention, we omit the outermost universal quantifier, and use **restricted quantifiers** such as $\exists x \in a$, $\forall x \in a$.

Axiom 1 (axiom of extensionality):

$$\forall x(x \in a \equiv x \in b) \rightarrow a = b.$$

This asserts that sets formed by the same elements are equal. The formula $x \in a(x \in b)$ is denoted by $a \subset b$. This means "a is a subset of b." Then Axiom 1 can be expressed by

$$a \subset b \wedge b \subset a \rightarrow a = b.$$

Axiom 2 (axiom of the unordered pair):

$$\exists x \forall y(y \in x \equiv y = a \vee y = b).$$

This asserts the existence, for any sets a and b, of a set x having a and b as its only elements. This x is called the **unordered pair** of a and b and is denoted by $\{a, b\}$. The notion of **ordered pair** is characterized by

$$(a, b) = (c, d) \equiv a = c \wedge b = d.$$

An example of such a set is $(a, b) = \{\{a, a\}, \{a, b\}\}$.

Axiom 3 (axiom of the sum set):

$$\exists x \forall y(y \in x \equiv \exists z \in a(y \in z)).$$

This asserts the existence for any set a of the sum (or union) x of all the sets that are elements of a. This x is denoted by Ua or $S(a)$. We write $a \cup b$ for $U\{a, b\}$ and a' for $a \cup \{a, a\}$.

Axiom 4 (axiom of the power set):

$$\forall x \exists y(y \in x \equiv \forall z \in y(z \in a)).$$

This asserts the existence for any set a of the power set x consisting of all the subsets of a. This x is denoted by $P(a)$. We have $S(P(a)) = a$, so S is a left inverse of P, and the products SP and PS are idempotent.

Axiom 5 (axiom of the empty set):

$$\exists x \forall y(\neg y \in x).$$

This asserts the existence of the empty set. The empty set is denoted by \emptyset or 0.

Axiom 6 (axiom of infinity):

$$\exists x(0 \in x \vee \forall y \in x(y' \in x)).$$

This asserts the existence of a set containing all the natural numbers, where $0, 1 = 0' = \{0\}$, $2 = 1' = \{0, 1\}$, $3 = 2' = \{0, 1, 2\}$. This definition of natural number is due to von Neumann.

Axiom 7 (axiom of separation):

$$\exists x \forall y(y \in x \equiv y \in a \wedge A(y)).$$

This asserts that the existence for any set a and a formula $A(y)$ of a set x consisting of all elements of a satisfying $A(y)$. This x is denoted by $\{y \in a \mid A(y)\}$. This is rather a schema for an infinite number of axioms, for there are an infinite number of $A(y)$. This axiom, also called the **axiom of comprehension** or **axiom of subsets**, was introduced by Zermelo.

For example, the set of all natural numbers is introduced by

$$\{x \in a \mid \forall y(0 \in y \wedge \forall z \in y(z' \in y) \rightarrow x \in y)\},$$

where a is a set satisfying Axiom 6. The set of all natural numbers is denoted by ω or N.

Axiom 8 (axiom of replacement):

$$\exists x \forall y \in a(\exists z A(y, z) \rightarrow \exists z \in x A(y, z)).$$

This asserts the existence for any set a of a set x such that for any y of a, if there exists a z satisfying $A(y, z)$ then such z exists in x. If the relation $A(y, z)$ determines a function, then the image of a set by the relation is included in a set, so by Axiom 7, the image is a set. This axiom was introduced by Fraenkel.

Axiom 9 (axiom of regularity):

$$\exists x A(x) \rightarrow \exists x(A(x) \wedge \forall y \in x(\neg A(y))).$$

This asserts that if there exists a set satisfying $A(x)$, then there is a set x satisfying $A(x)$ but every element y of x does not satisfy the property $A(y)$. This axiom is also called the **axiom of foundation**. Its dual expression, called \in -induction, is of course equivalent to Axiom 9.

Axiom 9' (axiom of E-induction):

$$\forall x(\forall y \in x A(y) \rightarrow A(x)) \rightarrow \forall x A(x).$$

Using this we can show that $\neg x \in x$, $\neg(x \in y \wedge y \in x)$, etc. If we assume the axiom of choice stated below, then this is equivalent to the nonexistence of an infinite descending sequence

$$x_n \in \dots \in x_2 \in x_1.$$

If a model of set theory satisfies the axiom of regularity and has an infinite descending sequence that is not in the model, then such a model is called a nonstandard model.

Axiom 10 (axiom of choice):

$$\forall x \in a \exists y A(x, y) \rightarrow \exists y \forall x \in a A(x, y(x)).$$

This asserts that if for any element x of a there is a set y such that $A(x, y)$, then there is a **choice function** y for the formula, i.e., $A(x, y(x))$ for all x in a . Usually a function is represented by its graph. A set f is called a function defined on a if

$$\forall x \forall y ((x, y) \in f \rightarrow x \in a), \quad \forall x \in a \exists y ((x, y) \in f),$$

$$\forall x \in a \forall y \forall z ((x, y) \in f \wedge (x, z) \in f \rightarrow y = z).$$

This formula is denoted by $\text{Fnc}(f)$; then the formula $A(x, f(x))$ is an abbreviation of

$$\text{Fnc}(f) \wedge \exists y ((x, y) \in f \wedge A(x, y)).$$

The axiom of choice is equivalent to many properties, such as the well-ordering theorem, Zorn's lemma, and Tikhonov's theorem on the product of compact spaces, and it is used widely and essentially in mathematics.

The system of axioms 1 to 9 is called **Zermelo-Fraenkel set theory** and is denoted by **ZF**; the system ZF minus the axiom of replacement is called **Zermelo set theory**, denoted by **Z**; and the system ZF plus the axiom of choice is denoted by **ZFC**.

The system **Z** is weaker than **ZF**. Indeed, the existence of the set ω of all the natural numbers and of $P(\omega)$, $P(P(\omega))$, can be proved in **Z**, but the existence of the set $\{\omega, P(\omega), P(P(\omega)), \dots\}$ cannot be proved in **Z**. However, we can prove its existence in **ZF**.

The theory **ZF** minus the axiom of infinity is called **general set theory**. Its consistency can be reduced to the consistency of the theory of natural numbers as follows. Let n be any natural number and $n = 2^{n_1} + 2^{n_2} + \dots + 2^{n_k}$ ($n_1 < n_2 < \dots < n_k$) be its binary expansion. We can define a model of general set theory by identifying a set with the natural number n and defining $m \in n$ by "m appears as one of the n_i ." In general set theory we can show the existence of 0 , $P(0)$, $P(P(0))$, but the existence of the set $\{0, P(0), P(P(0)), \dots\}$ cannot be proved. However, we can prove its existence in **Z**. The set consisting of all **hereditary** finite sets is the smallest model of general set theory.

C. Bernays-Gödel Set Theory

The existence of $\{u \mid A(u)\}$ for an arbitrary set-theoretic formula $A(u)$ cannot be deduced from the axioms of **ZF**. We call this object a **class** to distinguish it from sets. We introduce a generalized logical system of first-order predicate logic by adding to the logical system used in **ZF** class variables, class constants, and

inference rules with respect to quantifiers for classes. For any set-theoretic formula $A(u)$ in which no ϵ -bound class variable occurs, we adopt

$$\exists X \forall u (u \in X \leftrightarrow A(u))$$

as an axiom, where capital letters X are class variables. The set theory thus obtained is equivalent to **Gödel set theory** [11]. Von Neumann axiomatized set theory by making use of the notion of functions instead of that of classes [30]. In refining this theory and introducing the notion of classes, Bernays and Gödel [2, 11] initiated **Bernays-Gödel set theory**, **BG**.

ZF and **BG** are related as follows: Any formula provable in **ZF** is provable in **BG**, and any set-theoretic formula provable in **BG** and having neither class variable nor class constant is provable in **ZF**. In this sense, the systems can be regarded as essentially equivalent, but as **BG** has class variables and class constants, it is more convenient for expressing set-theoretic notions.

Von Neumann defined the following function R by ϵ -transfinite induction:

$$R(0) = \emptyset, \quad R(\alpha) = \bigcup_{\beta < \alpha} \mathfrak{P}(R(\beta)),$$

where α and β are ordinal numbers and $\bigcup_{\beta < \alpha} \mathfrak{P}(R(\beta))$ denotes the set sum of $\mathfrak{P}(R(0))$, $\mathfrak{P}(R(1))$, \dots , $\mathfrak{P}(R(\beta))$, \dots ($\beta < \alpha$). The function R can be defined by a set-theoretic formula, so it exists as a class. Now consider the model $M = M(\alpha)$ for a fixed ordinal number α . We define sets of the model M as elements of $R(\alpha)$, and classes of the model as subsets of $R(\alpha)$. We denote the \in relation of the model by \in_M . For classes X and Y of the model, we write $X \in Y$, $Y \not\in X$ if $X \in Y$. Then a necessary and sufficient condition for $R(\alpha)$ to be a model of **BG** is that α is an ϵ -inaccessible ordinal number ($\rightarrow 3$ 12 Ordinal numbers). The existence of an ϵ -inaccessible ordinal number cannot be deduced from the axioms of **ZF**. There is a series of studies of axiomatization of set theory in which any number of inaccessible ordinal numbers is assumed to exist [23]. When $R(\alpha)$ is a model of **BG** (**ZF**), it is called a **natural model** of **BG** (**ZF**). Furthermore, if H is $\bigcup_{\alpha} R(\alpha)$ then H satisfies all the axioms of **BG** (**ZF**). As we do not need the axiom of regularity for defining the class H , we see that **BG** (**ZF**) is consistent as long as **BG** (**ZF**) without the axiom of regularity is consistent.

D. Independence of the Continuum Hypothesis and the Axiom of Choice

These axiomatizations of set theory motivated a series of studies from the standpoint of the

foundations of mathematics on problems that remained unsolved after the appearance of Cantor's primitive set theory. Among these, the problem of the relation between the continuum hypothesis and the axiom of choice was central.

Consistency of the Axiom of Choice and the Continuum Hypothesis. Gödel [11] proved that if ZF without the axiom of choice is consistent, then the system obtained by adding to ZF the axiom of choice and the continuum hypothesis is also consistent. To show this, he constructed a model of ZF satisfying the axiom of choice and the generalized continuum hypothesis as follows. Assume first that M is an arbitrary domain of objects among which the \in relation is defined. By a formula on M , we understand a formula having constants of M as its only constants, having \in as its sole predicate symbol, and further having exclusively variables whose ranges are restricted to M . Let us denote by M' the totality of subsets of M defined by a formula $A(x)$ on M . Now we put $M_0 = \{\emptyset\}$, $M_{\alpha+1} = M'_\alpha$, $M_\beta = \bigcup_{\alpha < \beta} M_\alpha$, if β is a limit ordinal number. We call x **constructible** if $x \in M_\alpha$ for some ordinal number α , assumed to be less than the first inaccessible ordinal number, if any. We denote the totality of constructible sets by L , and the totality of sets of ZF by V . We call the assertion $V = L$, that is, every set is constructible, the **axiom of constructibility**. If we add this axiom to ZF, the axiom of choice and the generalized continuum hypothesis become provable. On the other hand, if we regard elements of L as sets of the model and the original \in relation as the \in relation of the model, we have a model of ZF in which the axiom of constructibility holds.

Independence of the Axiom of Choice and the Continuum Hypothesis. Since the result of Gödel, attempts have been made to prove that the axiom of choice is independent of the other axioms. Fraenkel constructed a model of set theory without satisfying the axiom of choice, starting from a countable number of objects that are not sets. A. Mostowski constructed in ZF a model of set theory having objects that are not sets, and he proved that the model satisfies the axiom: Every set can be linearly ordered, but does not satisfy the axiom of choice [27]. E. Mendelson constructed a model of set theory that does not satisfy the axiom of choice by making use of an infinite descending chain $a, \ni a_2 \ni a_3 \ni \dots$ [25]. These models, however, do not satisfy all the axioms of ZF or of Zermelo set theory minus the axiom of choice, even though they satisfy most

of the axioms. Consequently, they are not sufficient for proving independence.

P. J. Cohen [3, 4] proved the following results in connection with the independence of the axiom of choice, the continuum hypothesis, and $V = L$: If ZF is consistent, each of the following conditions has a model. (1) The axiom of choice as well as the generalized continuum hypothesis holds, but there exists an ω satisfying $\omega \notin L$ and $\omega \subset \omega$. (2) $\mathfrak{P}(\omega)$ is not well-ordered. (3) The axiom of choice holds, but the continuum hypothesis does not hold. (4) $\mathfrak{P}(\mathfrak{P}(\omega))$ cannot be linearly ordered.

By (1), we see that $V = L$ is independent of the axiom of choice and the generalized continuum hypothesis: (2) shows the independence of the axiom of choice; and (3) shows that the continuum hypothesis is independent of the axiom of choice. In (4), $\mathfrak{P}(\mathfrak{P}(\omega))$ corresponds to the set F of all real-valued functions defined on the interval $[0, 1]$, and (4) shows that the proposition " F can be linearly ordered" is not deducible in ZF without the axiom of choice.

E. Some Recent Results

Boolean-Valued Set Theory. D. Scott and R. M. Solovay defined models of set theory in which the set-theoretic formula has values in a complete Boolean algebra. This viewpoint is motivated by P. J. Cohen's original notions of the forcing relation and the generic filter.

According to the relation between sets and their representing functions, a function $f: A \rightarrow B$ corresponds to a subset of A when B is a complete Boolean algebra. So by transfinite induction we put

$$V_\alpha^{(B)} = \bigcup_{\beta < \alpha} P^{(B)}(V_\beta^{(B)}),$$

where $P^{(B)}(A)$ is the Boolean-valued power set of A defined by

$$u \in P^{(B)}(A) \equiv u: A \rightarrow B.$$

Let $V^{(B)}$ be the union of all $V_\alpha^{(B)}$'s; then the truth values of the formulas $\llbracket u \in v \rrbracket$ and $\llbracket u = v \rrbracket$ are defined by transfinite induction as

$$\llbracket u \in v \rrbracket = \sum_{x \in \text{dom}(v)} v(x) \cdot \llbracket x = u \rrbracket,$$

$$\llbracket u = v \rrbracket$$

$$= \prod_{x \in \text{dom}(u)} u(x) \Rightarrow \llbracket x \in v \rrbracket \cdot \prod_{x \in \text{dom}(v)} v(x) \Rightarrow \llbracket x \in u \rrbracket,$$

where $a \Rightarrow b$ is an abbreviation for the element $-a + b$ in B .

In $V^{(B)}$ the following properties are satisfied:

- (i) The truth value of any formula provable from ZFC is 1, the largest element of B .
- (ii) Any complete homomorphism $h: B \rightarrow B_2$

can be extended to

$$h: V^{(B_1)} \rightarrow V^{(B_2)},$$

so that the commutative relation

$$h(\llbracket A(u_1, \dots, u_n) \rrbracket) = \llbracket A(h(u_1), \dots, h(u_n)) \rrbracket$$

is satisfied for every restricted formula; further, if h is an epimorphism then it is satisfied for every formula.

(iii) (maximum principle) For any formula $A(x)$ there is an element u of $V^{(B)}$ such that

$$\llbracket \exists x A(x) \rrbracket = \llbracket A(u) \rrbracket.$$

Moreover, if $\llbracket \exists x R(x) \rrbracket = 1$ then there an element u of $V^{(B)}$ such that $\llbracket R(u) \rrbracket = 1$ and

$$\llbracket \exists x (R(x) \wedge A(x)) \rrbracket = \llbracket A(u) \rrbracket.$$

Property (i) means that if $[A] < 1$, then A is not provable from ZFC. Many consistency results in set theory are obtained in this way by constructing special partial order structures, topological spaces, or complete Boolean algebras.

Axiom of Strong Infinity. The axiom of infinity, Axiom 6, of ZF asserts that there exists a set including the set ω of all natural numbers. The set ω is the smallest infinite cardinal number. In general, the axiom of strong infinity is stated in a form such that certain special properties of ω are satisfied for an uncountable cardinal number.

Though such an axiom asserts the existence of large cardinal numbers, their properties sometimes reflect, for example, those of the real numbers, or of the set of real numbers, etc.

We state here some typical examples:

(i) **Weakly inaccessible cardinal number.** This is a regular limit cardinal number, that is,

$$\text{cf}(\omega_\alpha) = \omega_\alpha, \quad \forall \beta < \alpha (\omega_{\beta+1} < \omega_\alpha).$$

(ii) **Strongly inaccessible cardinal number.**

$$\text{cf}(\omega_\alpha) = \omega_\alpha, \quad \forall \beta < \alpha (2^{\omega_\beta} < \omega_\alpha).$$

(iii) **Weakly compact cardinal number.** ω_α is uncountable and the space $X = 2^{\omega_\alpha}$ with $<_{\omega_\alpha}$ -topology is ω_α -compact.

(iv) **Measurable cardinal number.** ω_α is uncountable and there is a nonprincipal $<_{\omega_\alpha}$ -additive 2-valued measure $\mu: P(\omega_\alpha) \rightarrow \{0, 1\}$.

(v) **Strongly compact cardinal number.** ω_α is uncountable and any product of $2 = \{0, 1\}$ with $<_{\omega_\alpha}$ -topology is ω_α -compact.

The strength of these properties increases in the order presented here. Properties (i) to (iii) are compatible with the axiom of constructibility $V=L$. And (iv) implies $V \neq L$ while it is compatible with $V=L(\mu)$. But (v) implies $V \neq L(a)$ for any set a .

There are many other stronger axioms of

infinity, and the consistency proofs of these "axioms" within ZF or ZFC would involve essential difficulties. Inconsistency proofs of these properties, if they exist, would be very interesting.

F. Examples of Results

(1) Cardinality and Cofinality (\rightarrow 3 12 Ordinal Numbers). (i) (W. B. Easton [6]) Let \mathfrak{M} be a model of ZFC (the Zermelo-Frankel axioms plus the axiom of choice) in which the GCH (generalized continuum hypothesis) is valid, i.e., $\forall \alpha (2^{\omega_\alpha} = \omega_{\alpha+1})$, and let g be a function from ordinals to ordinals in \mathfrak{M} such that $\forall \alpha, \beta (\alpha < \beta \Rightarrow \omega_{g(\alpha)} \leq \omega_{g(\beta)})$ and $\forall \alpha (\omega_\alpha < \text{cf}(\omega_{g(\alpha)}))$. Then there is a Boolean model \mathfrak{N} of ZFC, $\mathfrak{N} \supset \mathfrak{M}$, having the same cofinality and satisfying $2^{\omega_\alpha} = \omega_{g(\alpha)}$ for every regular cardinal. (This means that König's condition $\text{cf}(2^{\omega_\alpha}) > \omega_\alpha$ is the only restriction on the cardinality of powers of regular ω_α .)

(ii) (J. H. Silver [39]) Suppose

$$\omega < \text{cf}(\alpha) = \text{cf}(\omega_\alpha) < \omega_\alpha.$$

Then for any $\lambda < \text{cf}(\alpha)$

$$\forall \nu < \alpha (2^{\omega_\nu} \leq \omega_{\nu+\lambda}) \rightarrow 2^{\omega_\alpha} \leq \omega_{\alpha+\lambda}.$$

However, the validity of the implication

$$\forall n < \omega (2^{\omega_n} = \omega_{n+1}) \rightarrow 2^{\omega_\omega} = \omega_{\omega+1}$$

still remains an open question. And it is known that if there is a model for

$$\forall n < \omega (2^{\omega_n} = \omega_{n+1}), \quad 2^{\omega_\omega} > \omega_{\omega+1},$$

then there is a model in which there are many measurable cardinal numbers.

(iii) (K. Prikry [31]) Let ω_α be a measurable cardinal number in a model M of ZFC. Then there is a Boolean extension N of M in which the notion of cardinality is not changed but the notion of cofinality is changed, that is $\text{cf}(\omega_\alpha) = \omega_\alpha$ in M but $\text{cf}(\omega_\alpha) = \omega$ in N .

(iv) (Solovay) Let ω_α be a strongly compact cardinal number. Then for any strong-limit singular cardinal $\omega_\beta > \omega_\alpha$ the continuum hypothesis holds at ω_β ; that is, the implication

$$\text{cf}(\omega_\beta) < \omega_\beta, \quad \forall \gamma < \beta (2^{\omega_\gamma} < \omega_\beta) \rightarrow 2^{\omega_\beta} = \omega_{\beta+1}$$

is provable.

(2) Lebesgue Measurability and the Baire Property. As is well known, every Δ_1^1 (Borel) set (and consequently every Σ_1^1 (analytic) set) of real numbers is Lebesgue measurable and has the Baire property.

(i) (Gödel) $V=L$ implies the existence of a Δ_2^1 set of real numbers that is neither Lebesgue measurable nor has the Baire property.

Let DC denote the **principle of depending choice**:

$$\forall x \in a \exists y \in aP(x, y) \\ \rightarrow \exists f: \omega \rightarrow a \forall n \in \omega P(f(n), f(n+1)).$$

DC is adequate for the development of the classical notions of measure theory, such as Jordan decomposition, the Radon-Nikodym derivate, etc. Let I denote the hypothesis $\exists \alpha (\text{cf}(\omega_\alpha) = \omega_\alpha \wedge \forall \beta < \alpha (2^{\omega_\beta} < \omega_\alpha))$ (strongly inaccessible).

(i) (Solovay [40]) The consistency of ZFC and I implies that of ZF and either of the following two axioms:

(a) DC plus the hypothesis that every set of real numbers is Lebesgue measurable and has the Baire property.

(b) The axiom of choice plus the hypothesis that every set of real numbers definable by an ω sequence of ordinals is Lebesgue measurable and has the Baire property.

Axiom (a) means that DC is not strong enough to construct a (Lebesgue) nonmeasurable set, while Axiom (b) implies that every projective set is Lebesgue measurable and has the Baire property, and hence implies that the set of all constructible real numbers, as a Σ_2^1 set, has Lebesgue measure 0, and is of the first category.

(3) Martin's Axiom. Let B be a Boolean algebra. We say that B satisfies ω_α c.c. (chain condition) if the cardinality of every disjoint family of positive elements of B is at most ω_α . Let B^* denote the topological space consisting of all homomorphisms $h: B \rightarrow 2 = \{0, 1\}$ with the open base $U(a) = \{h \mid h(a) = 1\}$ ($a \in B$); then B^* is a * Baire space. Then **Martin's axiom** (MA) is: Let B be an ω c.c. Boolean algebra and $\alpha < 2''$. Then the intersection of a dense open sets is dense in B^* . Since $\{h \mid \sum_\nu h(a_\nu) = h(b)\}$ is dense and open in B^* if $\sum_\nu a_\nu = b$, MA means the existence of an $h \in B^*$ preserving any given set of α ($< 2''$) equations in B . If $2'' = \omega_1$, MA merely reduces to the Baire property of B^* . However, if $2'' > \omega_1$, then the ω c.c. hypothesis is essential, for there exists a B satisfying ω_1 c.c. such that B^* contains ω_1 dense open sets with empty intersection.

(i) (Solovay and S. Tennenbaum [43]) The consistency of ZF implies that of ZFC, MA, and $2^\omega > \omega_1$.

(ii) (D. A. Martin and Solovay [24]) ZFC, MA, and $2'' > \omega_1$ imply the propositions:

- (a) $\forall \omega_\alpha < 2^\omega (2^{\omega_\alpha} = 2^\omega)$, hence $2'' = 2^\omega$;
- (b) the totality of the first category sets of Lebesgue measure zero sets is α -additive for any $\alpha < 2''$;
- (c) every Σ_1^1 set of real numbers is Lebesgue measurable and has the Baire property.

(4) Suslin's Hypothesis (SH) is: Every dense, linear, order complete set without end points, having at most ω disjoint intervals, is order isomorphic to the continuum of real numbers.

(i) (T. J. Jech, Tennenbaum) The consistency of ZF implies that of (a) ZFC, \neg SH, and GCH, as well as (b) ZFC, \neg SH, and $2'' > \omega_1$

(ii) (R. Jensen [15]) The consistency of ZF implies that of ZFC, SH, and $2'' = \omega_1$.

(iii) (Solovay and Tennenbaum [43]) ZFC, MA, and $2'' > \omega_1$ imply SH.

(iv) (Jensen [15]) $V = L$ implies \neg SH.

(5) Measurable and Real-Valued Measurable Cardinals. A cardinal $\kappa > \omega$ is said to be **measurable** if there is a measure $\mu: \mathfrak{P}(\kappa) \rightarrow \{0, 1\}$ with (a) $\mu(\kappa) = 1$, (b) $\forall v < \kappa (\mu(\{v\}) = 0)$ and (c) $\mu(\sum_{v < \alpha} A_v) = \sum_{v < \alpha} \mu(A_v)$; and κ is said to be **real-valued measurable** if there exists a $\mu: \mathfrak{P}(\kappa) \rightarrow [0, 1]$ that satisfies (a), (b) and (c), and is not measurable.

Let MC (RMC) denote the existence of a measurable cardinal (real-valued measurable cardinal).

(i) (S. Ulam [45]) The existence of an ω -additive measure $\mu: \mathfrak{P}(A) \rightarrow [0, 1]$, with $\mu(A) = 1$ and $\forall x \in A (\mu(\{x\}) = 0)$, implies RMC or MC; RMC implies the existence of an extension of Lebesgue measure defined on $\mathfrak{P}([0, 1])$; every real-valued measurable cardinal is $\leq 2''$ and weakly inaccessible.

(ii) Every measurable cardinal is strongly (hyper) inaccessible, and (a) $\exists \alpha F(\alpha) \rightarrow \exists \alpha < \omega_1 F(\alpha)$ for any Σ_1^1 formula $F(\alpha)$ on the ordinal numbers, (b) $\exists \alpha F(\alpha) \rightarrow \exists \alpha < \mu_1 F(\alpha)$ for any Π_1^1 formula $F(\alpha)$ and the smallest measurable μ_1 . (Many results have been obtained concerning the ordinal magnitude of ω_1 and the measurable cardinals.)

(iii) (Solovay [41]) The consistency of ZFC and MC is equivalent to that of ZFC and RMC.

(iv) (Martin and Solovay [24]) ZFC, RMC, and MA are not consistent.

(v) (Lévy and Solovay [21]) The consistency of ZFC and MC implies that of ZFC, MC, MA, and $2^\omega > \omega_2$.

(vi) (Solovay) ZFC and MC imply that every Σ_2^1 set of real numbers is Lebesgue measurable.

(vii) (Martin, Solovay) ZFC, MC, MA, and $2'' > \omega_2$ imply that every Σ_3^1 set of real numbers is Lebesgue measurable and has the Baire property.

(viii) (Silver [38]) The consistency of ZFC and MC implies that of ZFC and MC as well as the existence of (Lebesgue) non-measurable Λ_3^1 sets of real numbers.

(6) Axiom of Determinateness. We consider the following infinite game. Let A be a set consist-

ing of functions ω to ω . The players I and II choose alternatively natural numbers n_i and m_i , and, for the resulting infinite sequence, if $(n_1, m_1, n_2, \dots, n_k, m_k, \dots) \in A$,

then I wins, otherwise II wins.

A function $\sigma: \bigcup_{n \in \omega} \omega^n \rightarrow \omega$ is called a **strategy**. If II plays according to the sequence $\underline{m} = (m_1, m_2, \dots, m_k, \dots)$ and I plays according to the strategy σ , then the resulting sequence is

$$\sigma^* \underline{m} = (\sigma(\emptyset), m_1, \sigma(m_1), \dots, m_k, \sigma(m_1, \dots, m_k), \dots).$$

The strategy σ is called a **winning strategy** for I of the game defined above, if for any play \underline{m} of II, the resulting play $\sigma^* \underline{m}$ is in the set A , and similarly for the winning strategy of II.

The axiom of determinateness AD is the assertion that for any $A \subset \omega^\omega$, the player I or II has a winning strategy.

(i) (J. Mycielski [28]) Assume ZF and AD; then we have:

- (a) axiom of choice for countable sequence of sets of real numbers;
- (b) every set of real numbers is Lebesgue measurable and has the Baire property, so the axiom of choice does not hold;
- (c) every uncountable set of real numbers contains a perfect set.

(ii) (Solovay) Assume ZF and AD; then we have:

- (a) cardinalities of ω_1 and 2^ω are not comparable;
- (b) ω_1 and ω_2 are measurable cardinals, but ω_3 is a singular cardinal such that $\text{cf}(\omega_3) = \omega_2$.

(iii) (Martin) In ZFC, every Borel, namely Δ_1^1 , game is determined.

(iv) (Martin, Solovay) In ZFC:

- (a) if there exists a measurable cardinal, then every analytic, namely Π_1^1 , game is determined;
- (b) if every Δ_2^1 -game is determined, then there is a model of ZFC with many measurable cardinal numbers.

Since the axiom of determinateness AD is very strong, there is some suspicion that it might contradict ZF. But at present there is no such evidence.

References

- [1] J. Bell, Boolean-valued models and independence proofs in set theory, Clarendon Press, 1977.
- [2] P. Bernays, A system of axiomatic set theory I, J. Symbolic Logic, 2 (1937), 655-77.
- [3] P. J. Cohen, The independence of the continuum hypothesis I, II, Proc. Nat. Acad. Sci. US, 50 (1963), 1143-1148; 51 (1964), 105-110.

- [4] P. J. Cohen, Set theory and the continuum hypothesis, Benjamin, 1966.
- [5] F. R. Drake, Set theory, North-Holland, 1974.
- [6] W. B. Easton, Power of regular cardinals, Ann. Math. Logic, 1 (1970), 139-178.
- [7] G. Föder, On stationary sets and regressive functions, Acta Sci. Math., 27 (1966), 1055-110.
- [S] A. Fraenkel, Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre, Math. Ann., (1922), 230-237.
- [9] H. Gaifman, Uniform extension operators for models and their applications, Sets, Models and Recursion Theory, North-Holland, 1967, 122-155.
- [10] F. Galvin and A. Haynal, Inequalities for cardinal powers, Ann. Math., (2) 101 (1975), 491-498.
- [11] K. Gödel, The consistency of the axiom of choice and of the generalized continuum hypothesis with the axioms of set theory, Ann. Math. Studies, Princeton Univ. Press, 1940.
- [12] K. Gödel, What is Cantor's continuum problem?, Amer. Math. Monthly, 54 (1947), 515-525.
- [13] T. Jech, Non-provability of Suslin's hypothesis, Comm. Math. Univ. Carolinae, 8 (1972), 229-308.
- [14] T. Jech, Set theory, Academic Press, 1978.
- [15] R. B. Jensen, The fine structure of the constructible hierarchy, Ann. Math. Logic, 4 (1972), 229-308.
- [16] A. Kanamori and M. Magidor, The evolution of large cardinal axioms in set theory, Lecture notes in math. 669, Springer, 1978, 99-275.
- [17] H. J. Keisler and A. Tarski, From accessible to inaccessible cardinals, Fund. Math., 53 (1965), 225-308; 57 (1965), 119.
- [18] K. Kunen, Inaccessibility properties of cardinals, Thesis, Stanford Univ.
- [19] K. Kunen and J. B. Paris, Boolean extensions and measurable cardinals, Ann. Math. Logic, 2 (1971), 359-378.
- [20] A. Lévy, Basic set theory, Perspectives in mathematical logic, Springer, 1979.
- [21] A. Lévy, and R. M. Solovay, Measurable cardinals and the continuum hypothesis, Israel J. Math., 5 (1967), 234-248.
- [22] M. Magidor, How large is the first strongly compact cardinal?, Ann. Math. Logic, 10 (1976), 33-57.
- [23] P. Mahlo, Zur Theorie und Anwendung der ρ_0 -Zahlen, Ber. Verh. Königl. Säch. Ges. Wiss. Leipzig. Math. Phys. Kl., 63 (1912), 108-112; 65 (1913), 268-282.
- [24] D. A. Martin and R. M. Solovay, Internal Cohen extensions, Ann. Math. Logic, 2 (1970), 143-178.
- [25] E. Mendelson, The independence of weak

- axiom of choice, *J. Symbolic Logic*, 21 (1956), 350-366.
- [26] Y. N. Moschovakis, Determinacy and prewellordering of the continuum, *Mathematical logic and foundations of set theory*, North-Holland, 1970, 244-262.
- [27] A. Mostowski, Über die Unabhängigkeit des Wohlordnungssatzes vom Ordnungsprinzip, *Fund. Math.*, 32 (1939), 201-252.
- [28] J. Mycielski, On the axiom of determinateness, *Fund. Math.*, 53 (1964), 205-224.
- [29] K. Namba, An axiom of strong infinity and analytic hierarchy of ordinal numbers, *Proc. Symp. Pure Math.*, 13 (1971), 279-319.
- [30] J. v. Neumann, Die Axiomatisierung der Mengenlehre, *Math. Z.*, 27 (1928), 699-752.
- [31] K. L. Prikry, Changing measurable into accessible cardinals, *Diss. Math.*, 68 (1970), 5-52.
- [32] J. B. Rosser, *Boolean-valued models for set theory*, Academic Press, 1969.
- [33] G. E. Sacks, Forcing with perfect closed sets, *Proc. Symp. Pure Math.*, 13 (1971), 331-355.
- [34] D. S. Scott, Measurable cardinals and constructible sets, *Bull. Acad. Polon. Sci., Sér. Math. Astr. Phys.*, 7 (1961), 145-149.
- [35] D. S. Scott and R. M. Solovay, *Boolean-valued models for set theory*, Lecture notes, Summer Institute on Axiomatic Set Theory, UCLA, 1967.
- [36] S. Shelah, A compactness theorem for singular cardinals, free algebras, Whitehead problem and transversals, *Israel J. Math.*, 21 (1975), 319-349.
- [37] J. H. Silver, Some applications of model theory in set theory, *Ann. Math. Logic*, 3 (1971), 45-110.
- [38] J. H. Silver, The consistency of the GCH with the existence of a measurable cardinal, *Proc. Symp. Pure Math.*, 13 (1971), 391-396.
- [39] J. H. Silver, On the singular cardinal problem, *Proc. Int. Congr. Math.*, Vancouver, 1 (1974), 265-286.
- [40] R. M. Solovay, A model of set theory in which every set of reals is Lebesgue measurable, *Ann. Math.*, (2) 92 (1970), 1-56.
- [41] R. M. Solovay, Real-valued measurable cardinals, *Proc. Symp. Pure Math.*, 13 (1971), 497-528.
- [42] R. M. Solovay, Strongly compact cardinal and the G.C.H., *Proc. Symp. Pure Math.*, 25 (1974), 365-372.
- [43] R. M. Solovay and S. Tennenbaum, Iterated Cohen extensions and Suslin's problem, *Ann. Math.*, (2) 94 (1971), 201-245.
- [44] G. Takeuti and W. B. Zaring, *Axiomatic set theory*, GTM 8, Springer, 1973.
- [45] S. Ulam, Zur Masstheorie in der allgemeinen Mengenlehre, *Fund. Math.*, 16 (1930), 140-150.

- [46] P. Vopenka and K. Hrbacek, On strongly measurable cardinals, *Bull. Acad. Polon. Sci., Sér. Sci. Math. Astr. Phys.*, 14 (1966), 587-591.
- [47] E. Zermelo, Untersuchungen über die Grundlagen der Mengenlehre I, *Math. Ann.*, 65 (1908), 261-281.

34 (11.5) Axiom of Choice and Equivalents

A. The Axiom of Choice

In set theory, the following axiom is known as the **axiom of choice**: For any nonempty family \mathfrak{A} of nonempty subsets of a set X , there exists a single-valued function f , called the **choice function**, whose domain is \mathfrak{A} , such that for every element A of \mathfrak{A} the value $f(A)$ is a member of A . This axiom is equivalent to each of the following three propositions: (1) If $\{A_\lambda\}_{\lambda \in \Lambda}$ is a family of sets not containing the empty set and with an index set Λ , then the Cartesian product $\prod_{\lambda \in \Lambda} A_\lambda$ is not the empty set. (2) If a set A is a disjoint union $\bigcup_{\lambda \in \Lambda} A_\lambda$ of a family of subsets $\{A_\lambda\}_{\lambda \in \Lambda}$, which does not contain the empty set, there exists a subset B , called the **choice set**, of A such that every intersection of B and A_λ ($\lambda \in \Lambda$) contains one and only one element. (3) For every mapping f from a set A onto a set B , there is a mapping from B to A such that $f \circ g = 1$, (identity mapping). Also equivalent to the axiom of choice are the well-ordering theorem and Zorn's lemma, which are discussed in the following sections.

B. The Well-Ordering Theorem

In 1904, E. Zermelo [2] first stated the axiom of choice and used it for his proof of the **well-ordering theorem**, which says that every set can be well-ordered by an appropriate ordering. Conversely, the well-ordering theorem implies the axiom of choice. Many important results in set theory can be obtained by using the axiom of choice, for example, that \aleph_1 cardinal numbers are comparable, or that various definitions of the finiteness or infiniteness of sets are equivalent. Various important theorems outside of set theory, e.g., the existence of bases in a \aleph_1 linear space, compactness of the direct product of compact \aleph_1 topological spaces (\aleph_1 -Tikhonov's theorem), the existence of a subset which is not \aleph_1 -Lebesgue measurable in Euclidean space, etc., are proved using the axiom of choice. But for those proofs the well-ordering theorem or Zorn's lemma (stated

below) are used more often than the axiom of choice.

Using the well-ordering theorem, it can be proved in the following manner that for every field k , any linear space X over k has a basis. Let $\{x_v\}_{v \in A}$ be an enumeration of X . By \uparrow transfinite induction, we can define the function f from X to $B = \{0, 1\}$ such that (i) if $x_0 = 0$, $f(x_0) = 0$; if $x_0 \neq 0$, $f(x_0) = 1$; (ii) for $v > 0$, if x_v is expressed as a linear combination of elements of $\{x_\mu \mid \mu < v, f(x_\mu) = 1\}$, $f(x_v) = 0$; otherwise $f(x_v) = 1$. Then $U = \{x \in X \mid f(x) = 1\}$ is a basis of X .

C. Zorn's Lemma

An ordered set X is called an **inductively ordered set** if every \uparrow totally ordered subset of X has an \uparrow upper bound. A condition C for sets is called a **condition of finite character** if a set X satisfies C if and only if every finite subset of X satisfies C . A condition C for functions is called a **condition of finite character** if C is a condition of finite character for the graph of the function. **Zorn's lemma** [4] can be stated in any one of the following ways, which are all equivalent to the axiom of choice. It is often more convenient to use than the axiom of choice or the well-ordering theorem.

- (1) Every inductively ordered set has at least one maximal element.
- (2) If every well-ordered subset of an ordered set M has an upper bound, then there is at least one maximal element in M .
- (3) Every ordered set A has a well-ordered subset W such that every upper bound of M belongs to W .
- (4) For a condition C of finite character for sets, every set X has a maximal (for the relation of the inclusion) subset of X that satisfies C .
- (5) Let C be a condition of finite character for functions from X to Y . Then, in the set of functions that satisfy C , there is a function whose domain is maximal (for the relation of the inclusion).

Using Zorn's lemma (1), we can prove again in the following way any linear space X over a field k has a basis. Let \mathfrak{A} be the set of all non-empty subsets A of X such that arbitrary finite subsets of A are linearly independent over k . \mathfrak{A} is not empty. If we order \mathfrak{A} by the relation of inclusion, then \mathfrak{A} is an inductively ordered set. By Zorn's lemma (1), there is a maximal element U of \mathfrak{A} . Since U is maximal, U is a basis of X .

The same theorem is proved as follows using Zorn's lemma (4). Condition C for the subset A of X , that arbitrary finite subsets of A are linearly independent over k , is a condition

of finite character. Hence there is a maximal subset U that satisfies C and is a basis of X .

Concerning recent developments regarding the axiomatic basis for the axiom of choice \rightarrow 33 Axiomatic Set Theory D.

References

- [1] G. Cantor, Über unendliche, lineare Punktmannigfaltigkeiten, Math. Ann., 21 (1883), 545-591. (Gesammelte Abhandlungen, Springer, 1932.)
- [2] E. Zermelo, Beweis, dass jede Menge wohlgeordnet werden kann, Math. Ann., 59 (1904), 514-516.
- [3] E. Zermelo, Neuer Beweis für die Möglichkeit einer Wohlordnung, Math. Ann., 65 (1908), 107-128.
- [4] M. Zorn, A remark on method in transfinite algebra, Bull. Amer. Math. Soc., 41 (1935), 667-670.
- [5] J. W. Tukey, Convergence and uniformity in topology, Ann. Math. Studies, Princeton Univ. Press, 1940.
- [6] N. Bourbaki, Eléments de mathématique I, Théorie des ensembles, ch. 3, Actualités Sci. Ind., 1243, Hermann, second edition, 1966; English translation, Theory of sets, Addison-Wesley, 1968.

35 (1.2) Axiom Systems

A. History

A mathematical theory is based on a specific system of axioms, i.e., a system of hypotheses from which the whole theory is deduced without reliance on other assumptions.

One of the first deductive methods of mathematical reasoning was utilized by Thales, who returned to Greece from Egypt with the knowledge of surveying methods, and who deduced additional results from that empirical knowledge. His method gave impetus to the development of Greek geometry, which flowered with the Pythagorean school and research by members of Plato's Academy. In the course of this development, the deductive method led to the idea of constructing the whole theory upon a system of "absolutely obvious" statements from which the whole theory could be deduced. Euclid systematized Greek geometry in his *Elements* utilizing this idea. His work became the basis of geometry after the Renaissance, and Greek geometry came to be called \uparrow Euclidean geometry. In the *Elements*

Euclid called the basic obvious statements **common notions** when they were of general nature, and **postulates** when they were specifically geometric. Both were later called **axioms** (or **postulates**).

Among the axioms stated by Euclid, the “fifth postulate” concerning parallels was longer and more complicated than the other axioms. Many efforts were made to deduce this particular axiom from the other axioms. The failure of these attempts suggested the possibility of establishing a †non-Euclidean geometry, which was actually done by N. I. Lobachevskii and J. Bolyai, who replaced the fifth postulate by its negation and showed that the new system of “axioms” was as valid as the classical one. This development naturally led to a new evaluation of the idea of axioms, and eventually the traditional concept of recognizing the axioms as obvious truths was replaced by the understanding that they are hypotheses for a theory. D. Hilbert [1] established the latter idea as **axiomatization** and claimed that the whole science of mathematics should be built upon a system of axioms. His idea became the foundation of present-day mathematics. Hilbert reorganized classical geometry based upon his idea and published his result in *Foundations of Geometry* (1899).

B. Systems of Axioms

The **system of axioms** of a theory, i.e., the system of basic hypotheses from which we hope to deduce the whole theory, is written in **undefined terms** (or in terms of **undefined concepts**) by means of which all other terms are defined. On the other hand, a given theory is **axiomatized** by specifying such a system of axioms upon which the theory may be reorganized. It should also be noted that a system of axioms determines a †structure (→ 409 Structures).

A system of axioms is considered to be mathematically valid if and only if it is **consistent**. It is also desirable that the axioms in such a system be mutually **independent** (i.e., the negation of any one of the axioms is still consistent with the others). When such a system is not independent, it can be simplified by deleting redundant axioms from it.

When any two models of a system of axioms are isomorphic to each other, we call the system **complete** or **categorical** (→ 409 Structures). For example, the system of axioms (I)–(V) postulated by Hilbert as the foundation for Euclidean geometry is **complete** (→ 155 Foundations of Geometry), whereas the systems of axioms for the theories of †groups, †rings, or †fields are not complete since there

are nonisomorphic groups, etc. Although it is desirable that the systems of axioms postulated for a given theory (e.g., the †theory of real numbers, or Euclidean geometry) be **complete**, the study of partial systems that may not be complete is also important (→ 156 Foundations of Mathematics).

References

- [1] D. Hilbert, *Axiomatisches Denken*, Math. Ann., 78 (1918), 405–415. (Gesammelte Abhandlungen III, Springer, 1935, 1466156.)
- [2] N. Bourbaki, *Eléments de mathématique*, Théorie des ensembles, ch. 1–4 and fascicule de résultats, *Actualités Sci. Ind.*, 1212c, 1243b, 1258a, 1141d, Hermann, 1960, 1967, 1966, 1964.

B

36 (XII.1 7) Banach Algebras

A. Definition [1-4]

A real or complex [†]Banach space R is called a **Banach algebra** (or **normed ring**) if a multiplication law between elements of R is introduced that makes R an [†]associative algebra and if $\|xy\| \leq \|x\| \cdot \|y\|$ is satisfied. The complex case is generally easier to handle, and a real Banach algebra can always be embedded in a complex one isomorphically and isometrically. When a Banach algebra contains a unity element e with respect to the multiplication, i.e., it is **unital**, we can suppose $\|e\| = 1$. If it is not unital, we can adjoin a unity element to it.

B. Examples

1. Let $C_0(\mathfrak{M})$ be the set of complex- (real-) valued continuous functions $x(\xi)$ on a locally compact space \mathfrak{M} that vanish at infinity. Endowed with the pointwise operations and the [†]supremum norm (- 168 Function Spaces), it forms a complex (real) Banach algebra. In the complex case, it has the involution $x^*(\xi) = \overline{x(\xi)}$.

2. Let X be a Banach space. The set $\mathcal{B}(X)$ of [†]bounded linear operators on X forms a Banach algebra if we define addition, multiplication by scalars, multiplication between elements, and the norm in the usual fashion.

3. Let G be a locally compact Hausdorff group (- 423 Topological Groups) and μ be its [†]left-invariant Haar measure. The Banach space $L_1(G)$ (with respect to μ) can be made into a Banach algebra by defining $xy(g) = \int x(h)y(h^{-1}g)d\mu(h)$ (- Section L).

4. Every subalgebra of a Banach algebra R , which is closed in the norm topology, is a Banach algebra with respect to the original algebraic operations and norm and is called a **closed subalgebra** of R .

Throughout this article, all Banach algebras are supposed to be complex Banach algebras unless otherwise specified.

C. Spectrum of an Element

We define a new operation $x \cdot y$ in R by setting $x \cdot y = x + y - xy$. If $x \cdot y = y \cdot x = 0$, y is called a **quasi-inverse** of x . When a unity element e exists, y is a quasi-inverse of x if and only if $e - y$ is an inverse of $e - x$. For a complex number λ such that $|\lambda| > \|x\|$, we see that

$\lambda^{-1}x$ possesses a quasi-inverse y given by the strongly convergent series $y = -\sum_{n=1}^{\infty} \lambda^{-n}x^n$. The set of all complex numbers λ such that $\lambda^{-1}x$ does not have a quasi-inverse is called the **spectrum** of x and is denoted by $\text{Sp}(x)$. (When x itself does not have an inverse, in particular, when R does not contain a unity element, we include 0 in this set.) $\text{Sp}(x)$ is a bounded closed set in the complex plane, and $\sup\{|\lambda| \mid \lambda \in \text{Sp}(x)\} = \lim_{n \rightarrow \infty} \|x^n\|^{1/n}$. In Example 2, this spectrum is the spectrum of the operator x , and the problem of determining the behavior of this spectrum constitutes one of the central problems in the theories of Banach and Hilbert spaces.

D. Representations of a General Banach Algebra [1, 2]

We understand by a **representation** of a Banach algebra R on a Banach space X any algebraic homomorphism π of R into the algebra $\mathcal{B}(X)$ (see Example 2) satisfying $\|\pi(x)\| \leq \|x\|$ for all $x \in R$. We call X the **representation space**. A Banach algebra always possesses an isomorphic and isometric representation. Especially important are the **irreducible representations**. A vector subspace (closed or not) Y of X is an invariant subspace if $\pi(x)Y \subset Y$ for any $x \in R$. A representation is called algebraically irreducible if the invariant subspaces are trivial, i.e., they are only $\{0\}$ or X . A representation is said to be topologically irreducible if closed invariant subspaces are trivial. The [†]kernel of an algebraically irreducible representation is called a **primitive ideal**, which can alternatively be defined in the following way: A left ideal $J (\neq \{0\}, R)$ is **regular**, by definition, if R contains an element u , a unity element modulo J , such that $x - xu \in J$ for any $x \in R$. Such an ideal is always contained in a maximal left ideal, which in turn is necessarily regular and closed. A two-sided ideal I is primitive if it is the set of elements a in R for which $aR \subset J$, where J is some fixed regular maximal left ideal. If R is commutative, a primitive ideal is a regular maximal ideal, and conversely. The intersection of all primitive ideals is called the **radical** of R , and when it is $\{0\}$, R is called **semisimple**.

The set of primitive ideals \mathfrak{J} is known as the **structure space** of R , in which the **hull-kernel topology** (or **Jacobson topology**) is introduced. The [†]closure of a set \mathfrak{A} in \mathfrak{J} is, under this topology, the set of primitive ideals containing the intersection of the ideals in \mathfrak{A} . This topology is rather intractable; even in commutative cases, it does not coincide with the Gel'fand topology (\rightarrow Section E), in general.

E. The Gel'fand Representation of a Commutative Banach Algebra

A complex Banach algebra is a field if and only if it coincides with the field \mathbb{C} of complex numbers (**Gel'fand-Mazur theorem**). This is the most fundamental fact in the study of commutative Banach algebras. Now let R be a commutative Banach algebra. Then every regular maximal ideal J of R is closed and the quotient algebra R/J is isomorphic to \mathbb{C} , so that the quotient mapping $R \rightarrow R/J$ is viewed as a complex-valued homomorphism of R . Conversely, if φ is a nonzero complex-valued homomorphism of R , it has norm ≤ 1 as a linear functional on R and its kernel, say R_φ , is a regular maximal ideal of R such that φ is exactly the quotient mapping $R \rightarrow R/R_\varphi$. The set $\Pi(R)$ of nonzero complex-valued homomorphisms (or regular maximal ideals) of R endowed with the weak* topology of the Banach space dual R' of R is called the **maximal ideal space** of R . $\mathfrak{M}(R)$ is a locally compact Hausdorff space, and its topology is called the **Gel'fand topology**. For each $x \in R$ we define a function \hat{x} on $\mathfrak{M}(R)$ by setting $\hat{x}(\varphi) = \varphi(x)$. Then the mapping $x \rightarrow \hat{x}$ is a homomorphism of R into the algebra $C_0(\mathfrak{M}(R))$ of all continuous complex-valued functions on $\mathfrak{M}(R)$ vanishing at infinity. This is the **Gel'fand representation** of R whose image, the **Gel'fand transform** of R , is denoted by \hat{R} . Concerning this, we have (1) $\mathfrak{M}(R)$ is compact if R has a unity element; (2) $\text{Sp}(x)$ equals the closure of the range $\hat{x}(\mathfrak{M}(R))$ of \hat{x} ; (3) the representation $x \rightarrow \hat{x}$ is norm-decreasing if $C_0(\mathfrak{M}(R))$ is endowed with the supremum norm; (4) $\|\hat{x}\|_\infty (= \sup\{|\hat{x}(\varphi)| \mid \varphi \in \mathfrak{M}(R)\})$ equals $\lim_{n \rightarrow \infty} \|x^n\|^{1/n}$. The kernel of the Gel'fand representation of R is the **radical** of R , which consists of **generalized nilpotent elements** of R .

F. Banach Star Algebras

An **involution** in a Banach algebra R is an operation $x \rightarrow x^*$ that satisfies (1) $(x + y)^* = x^* + y^*$; (2) $(\lambda x)^* = \bar{\lambda} x^*$; (3) $(xy)^* = y^* x^*$; (4) $(x^*)^* = x$. A Banach algebra with an involution is called a **Banach *-algebra**. A ***-homomorphism** Φ between two Banach *-algebras is an algebraic homomorphism which preserves involutions, i.e., $\Phi(x^*) = \Phi(x)^*$. To represent a Banach *-algebra we prefer a ***-representation**, i.e., a representation $x \rightarrow T_x$ on a Hilbert space such that T_{x^*} is equal to the adjoint T_x^* of T_x for any $x \in R$.

G. C*-Algebras [5–10]

A Banach *-algebra A satisfying $\|x^* x\| = \|x\|^2$ for all $x \in A$ is called a **C*-algebra**. Every C*-

algebra is *-isomorphic and -isometric to a Banach algebra of operators on a Hilbert space (see Example 2) that contains, along with an operator, its adjoint (the **Gel'fand-Vaĭmark theorem**). A C*-algebra is semi-simple, and a commutative C*-algebra A is *-isomorphic to $C_0(\mathfrak{M}(A))$ under the Gel'fand representation. A topologically irreducible *-representation of a C*-algebra is also algebraically irreducible (R. V. Kadison), and the set of unitary equivalence classes of these irreducible *-representations is called the **dual space**. It becomes a topological space if we introduce the hull-kernel topology inherited from the structure space, but other topologies are also introduced. Moreover, for the study of separable C*-algebras, Borel structure in the sense of G. W. Mackey is a very powerful tool. C*-algebras also have intimate connections with the theory of unitary representations of a topological group (see below) and with quantum physics. Many works have been published on *-representations, dual spaces, etc.

A linear functional ω on a C*-algebra A is said to be **positive** if $\omega(x^* x) \geq 0$ for any $x \in A$. For any nonzero positive linear functional ω of A , there exist a *-representation (π_ω, H_ω) of A on a Hilbert space H , and a vector $\xi_\omega \in H_\omega$ such that $\omega(a) = (\pi_\omega(a)\xi_\omega, \xi_\omega)$ for all $a \in A$ and that the subspace $\{\pi_\omega(a)\xi_\omega \mid a \in A\}$ is dense in H_ω . We call (π_ω, H_ω) a **cyclic representation** of A induced by ω . A *-representation (π, H) of A is called **universal** if for any *-representation (ρ, K) of A there exists a σ -weakly continuous *-homomorphism $\tilde{\rho}$ of $n(A)''$ onto $p(A)''$ such that $\rho(a) = (\tilde{\rho} \circ \pi)(a)$ for all $a \in A$, where $\pi(A)''$ is the double commutant of $\pi(A)$ and so is a von Neumann algebra (\rightarrow 308 Operator Algebras). Von Neumann algebras $\pi(A)''$ for universal representations π of A are mutually *-isomorphic, so that they determine the **enveloping von Neumann algebra** of A . Especially, the direct sum π of all cyclic representations π_ω is a universal representation and there is a unique isometric isomorphism of $n(A)''$ onto the second dual A^{**} of A that is bicontinuous in the σ -weak topology of $n(A)''$ and the $\sigma(A^{**}, A^*)$ -topology of A^{**} . So A^{**} is identified with the enveloping von Neumann algebra of A .

H. Some Classes of C*-Algebras

Let A be a C*-algebra. If A is a **CCR (liminal C*-) algebra** if it is mapped to the algebra of compact operators under any irreducible *-representation. It is a **GCR (postliminal C*-) algebra** if every nonzero quotient C*-algebra of A has a nonzero CCR closed two-sided ideal. These classes of C*-algebras have inter-

esting properties. For example, the dual space of a separable C^* -algebra is a ${}^{\dagger}T_0$ space (${}^{\dagger}T_1$ space) if and only if the algebra is GCR (CCR) (J. Glimm).

A is called an **AF (approximately finite) algebra** if it is the uniform closure of an increasing sequence of finite-dimensional C^* -algebras (0. Bratteli). AF algebras are classified by so-called dimension groups. To see this, we call an ordered Abelian group G (written additively) a **Riesz group** if (a) for any integer $n > 0$ and any $g \in G$, $ng \geq 0$ implies $g \geq 0$ and (b) for any g_i, h_j ($1 \leq i \leq m, 1 \leq j \leq n$) in G with $g_i \leq h_j$ for all i and j , there exists $k \in G$ with $g_i \leq k \leq h_j$ for all i, j . Then the isomorphism classes of AF algebras correspond bijectively to the isomorphism classes, as local semi-groups, of generating upward-directed hereditary subsets of the positive cones of countable Riesz groups, and a dimension group, defined otherwise, is exactly a countable Riesz group (G. Elliot, E. G. Effros, D. E. Handelman and C. L. Shen) (\rightarrow [II]).

Now consider two C^* -algebras A and B . The algebraic † tensor product $A \otimes B$ over C becomes a $*$ -algebra in the natural way. A norm β on $A \otimes B$ is a **C^* -crossnorm** if $\|xy\|_{\beta} \leq \|x\|_{\beta} \|y\|_{\beta}$ and $\|x^*x\|_{\beta} = \|x\|_{\beta}^2$ for $x, y \in A \otimes B$ and $\|a \otimes b\|_{\beta} \leq \|a\| \|b\|$ for $a \in A, b \in B$. The completion of $A \otimes B$ under such a norm β is a C^* -algebra, which is denoted by $A \otimes_{\beta} B$. There are two special C^* -crossnorms on $A \otimes B$: $\|\cdot\|_{\min}$ and $\|\cdot\|_{\max}$. The former is called the spatial (minimal, injective) C^* -crossnorm and is defined by $\|x\|_{\min} = \sup_{\varphi, \psi} \|(\varphi \otimes \psi)(x)\|$, where φ and ψ run over all $*$ -representations of A and B , respectively. The algebra $A \otimes_{\min} B$ is called the **spatial tensor product** of A and B . The latter is called the greatest C^* -crossnorm and is defined by $\|x\|_{\max} = \sup_T \|T(x)\|$, where T runs over all $*$ -representations of $A \otimes B$. And the algebra $A \otimes_{\max} B$ is called the **projective C^* -tensor product** of A and B . Any C^* -crossnorm β on $A \otimes B$ lies in between these two norms, i.e., $\|x\|_{\min} \leq \|x\|_{\beta} \leq \|x\|_{\max}$ for $x \in A \otimes B$. A C^* -algebra A is called **nuclear** if $A \otimes B$ has a unique C^* -crossnorm for any C^* -algebra B . It is known that any GCR algebra is nuclear and that an inductive limit of nuclear C^* -algebras (e.g., any AF algebra) is nuclear. Given a linear mapping φ between C^* -algebras A and B , we define for any integer $n \geq 1$ a linear mapping $\varphi_n: A \otimes M_n \rightarrow B \otimes M_n$ by setting $\varphi_n(x \otimes e_{ij}) = \varphi(x) \otimes e_{ij}$, where $\{e_{ij}\}$ are the matrix units for the C^* -algebra M_n of $n \times n$ complex matrices. φ is said to be **completely positive** if φ_n is positive for any $n \geq 1$. A C^* -algebra A is then called **injective** if, for any C^* -algebras B, C with $B \subseteq C$ and any completely positive contraction $\varphi: B \rightarrow A$, there

exists a completely positive contraction $\tilde{\varphi}: C \rightarrow A$ that extends φ . A C^* -algebra A is nuclear if and only if its enveloping von Neumann algebra A^{**} is injective. Every injective C^* -algebra B is not necessarily a von Neumann algebra but every **AW*-algebra**, i.e., every maximal commutative $*$ -subalgebra M of B , is monotone complete (i.e., every increasing net of self-adjoint elements has a least upper bound in M) and the lattice of projections in B is conditionally complete.

I. Crossed Product

Let A be a C^* -algebra, G a locally compact Hausdorff group with left-invariant measure μ , and α a continuous action of G on A as $*$ -automorphisms $\alpha_g, g \in G$. Let $K(G, A)$ be the linear space of continuous functions from G to A with compact support, which is a $*$ -algebra under the multiplication $y * z$ and the involution y^* given by $(y * z)(g) = \int_G y(h) \alpha_h(z(h^{-1}g)) d\mu(h)$ and $y^*(g) = \Delta(g^{-1}) \cdot \alpha_g(y(g^{-1})^*)$ for $y, z \in K(G, A)$, where Δ is the † modular function of G . By completing $K(G, A)$ under the norm $\|y\|_1 = \int_G \|y\| d\mu(g)$, we get a Banach $*$ -algebra $L_1(G, A)$. $L_1(G, A)$ is then made into a C^* -algebra by furnishing the norm $\|x\| = \sup_{\pi} \|\pi(x)\|$, where π runs over all $*$ -representations of $L_1(G, A)$. This C^* -algebra, denoted by $A \otimes_{\alpha} G$, is the **crossed product** of A by G relative to the action α . Crossed products are useful in the structure theory of C^* -algebras.

J. Extensions by C^* -Algebras (BDF Theory) [12]

Let H be a separable infinite-dimensional Hilbert space and $\mathcal{L}\mathcal{C}(H)$ the ideal of $\mathcal{B}(H)$ consisting of all compact operators. The quotient C^* -algebra $\mathcal{B}(H)/\mathcal{L}\mathcal{C}(H)$ is called the **Calkin algebra**. We denote it by $\mathcal{Q}(H)$ and the quotient mapping by $\pi: \mathcal{B}(H) \rightarrow \mathcal{Q}(H)$. By an **extension** of $\mathcal{L}\mathcal{C}(H)$ by a separable unital C^* -algebra A we mean a unital (preserving unity elements) $*$ -isomorphism τ of A into $\mathcal{Q}(H)$. We call two extensions $\tau_1, \tau_2: A \rightarrow \mathcal{Q}(H)$ **equivalent** if, for some unitary operator $u \in \mathcal{B}(H)$, $\pi(u)\tau_1(a)\pi(u)^* = \tau_2(a)$ for any $a \in A$, and denote by $\text{Ext}(A)$ the set of all equivalence classes $[t]$ of extensions τ by A . $\text{Ext}(A)$ forms a commutative semigroup with respect to the addition $[\tau] = [\tau_1] + [\tau_2]$, where $\tau(a) = \tau_1(a) \oplus \tau_2(a) \in \mathcal{Q}(H) \oplus \mathcal{Q}(H) \subset \mathcal{Q}(H \oplus H)$. Call an extension $\tau: A \rightarrow \mathcal{Q}(H)$ trivial if there is a unital $*$ -isomorphism σ of A into $\mathcal{B}(H)$ with $\tau = \pi\sigma$. Then (1) all trivial extensions are equivalent,

and we define the unity element of the semi-group $\text{Ext}(A)$ (D. Voiculescu); (2) $\text{Ext}(A)$ is a group if A is commutative (L. G. Brown, R. G. Douglas, and P. A. Fillmore) or, more generally, if A is nuclear (M. D. Choi and E. G. Effros); (3) $\text{Ext}(A)$ is not always a group (J. Anderson). For further information \rightarrow 390 Spectral Analysis of Operators J.

K. Derivations in C^* -Algebras

A linear operator δ in a C^* -algebra A is a **derivation** in A if its domain $D(\delta)$ is a dense subalgebra of A and $\delta(xy) = \delta(x)y + x\delta(y)$ for $x, y \in D(\delta)$. δ is a **$*$ -derivation** if, moreover, $x \in D(\delta)$ implies $x^* \in D(\delta)$ and $\delta(x^*) = \delta(x)^*$. δ is called bounded (unbounded, closed) if it is bounded (unbounded, closed) as a linear operator.

Every bounded derivation δ is expressed as $\delta(x) = \tilde{a}x - x\tilde{a}$ with some \tilde{a} in the enveloping von Neumann algebra A^{**} of A (R. V. Kadison and S. Sakai). The element \tilde{a} can be taken from the **multiplier algebra** $M(A) = \{b \in A^{**} \mid bA + Ab \subseteq A\}$ of A if A is simple (Sakai) or if A has continuous trace (C. A. Akemann, Elliott, G. K. Pedersen, and J. Tomiyama). If A is separable, we see that all bounded derivations in A are given by elements in $M(A)$ if and only if A is the C^* -direct sum of a family of simple C^* -algebras and a C^* -algebra with continuous trace (Elliott, Akemann, and Pedersen) [13] and that every bounded derivation in the quotient C^* -algebra A/I , I being any closed two-sided ideal, can be lifted to a derivation in A (Pedersen).

Next we consider unbounded derivation. By a **$*$ -automorphism group** on A we mean a \dagger one-parameter group $\rho_t, t \in \mathbf{R}$, of $*$ -automorphisms of A such that, for each $x \in A$, $\rho_t x$ is continuous in $t \in \mathbf{R}$. A **C^* -dynamical system** is a pair consisting of a C^* -algebra A and a $*$ -automorphism group ρ_t on A . The fact that the time evolution of a physical system is often represented by such a dynamical system has made the study of unbounded derivations quite active. We have to see if a given derivation δ is \dagger closable and if, in case of a closable δ , the closure $\bar{\delta}$ generates a $*$ -automorphism group. Sample results: (1) If a $*$ -derivation δ is **well-behaved** in the sense that for every self-adjoint x in $D(\delta)$ there exists a \dagger state φ with $|\varphi(x)| = \|x\|$ and $\varphi(\delta(x)) = 0$, then δ is closable and its closure is well-behaved (A. Kishimoto and Sakai); (2) if a $*$ -derivation δ is closable, its closure $\bar{\delta}$ generates a $*$ -automorphism group if and only if δ is well-behaved and $(1 \pm \delta)D(\delta)$ is dense in A (R. T. Powers and S. Sakai; O. Bratteli and D. W. Robinson). For further results \rightarrow [14].

L. Applications to the Theory of Topological Groups [2, 3, 6, 15]

Banach algebras have many applications in different branches of mathematics and quantum physics. Here we mention some that concern topological groups. Let G be a locally compact Hausdorff group and μ be its left-invariant measure. We make the complex algebra $L_1(G)$ of Example 3 into a Banach **$*$ -algebra** by defining $x^*(g) = \overline{x(g^{-1})}\Delta(g^{-1})$, where Δ is the \dagger modular function of G . This is called the **L_1 -algebra** (or **group algebra**). It is not C^* but is semisimple. Considering a unitary representation of G is equivalent to considering a $*$ -representation of the L_1 -algebra. Replace the norm of $x \in L_1(G)$ by $\sup \|\pi(x)\|$, where the supremum is taken over all the $*$ -representations. The new norm satisfies the C^* condition, and the completion of $L_1(G)$ with respect to this norm is a C^* -algebra, which we call the **C^* -group algebra** of G . The dual of the C^* -group algebra thus defined is called the dual of the group G , and this notion plays an important role in the study of topological groups. Unitary representations of a group G , $*$ -representations of the L_1 -algebra of G , and $*$ -representations of the C^* -group algebra of G are all characterized by positive definite functions on G . A function $p(g)$ is **positive definite**, by definition, if it is measurable on G and

$$\iint p(g^{-1}h)x(g)x(h)d\mu(g)d\mu(h) \geq 0$$

for any continuous function $x(g)$ with compact support.

The Abelian case. When G is an Abelian group (\rightarrow 422 Topological Abelian Groups), a regular maximal ideal M of the L_1 -algebra R of G and a character γ of G are in a one-to-one correspondence by the relation

$$\hat{x}(M) = \int x(g)\gamma(g)d\mu(g)$$

(the left-hand side is the value of x at M under the Gel'fand representation). Moreover, the set of regular maximal ideals of R provided with the Gel'fand topology and the set \hat{G} of characters of G provided with the Pontryagin topology (the \dagger character group of G) are homeomorphic by this correspondence. Therefore the Gel'fand transform of an element x of the L_1 -algebra R is seen to be a function $\hat{x}(\gamma)$ on \hat{G} defined by the integral on the right-hand side in the above expression, which is properly called the **Fourier transform** of x . Of course, the Fourier transform can be defined for other classes of functions (e.g., for the L_2 -space over G , the Fourier transform in the sense of Plan-

cherel), and classical theories of Fourier series, Fourier integrals, and harmonic analysis (- 192 Harmonic Analysis) are studied from a more extensive point of view. Thus the statement that the Fourier transform of an element x of the L_1 -algebra of G is a continuous function vanishing at infinity (\rightarrow Section E), for example, is a version of the classical \dagger Riemann-Lebesgue theorem. **Bochner's theorem** in classical \dagger Fourier analysis is restated thus: A continuous \dagger positive definite function on an Abelian group can be put in the form

$$p(g) = \int \gamma(g) d\rho(\gamma),$$

where ρ is a uniquely determined bounded positive \dagger Radon measure on the character group \hat{G} . Developing these theories further, we obtain an alternative proof of the \dagger Pontryagin duality theorem (H. Cartan and R. Godement). A closed ideal I in the L_1 -algebra R determines a set $Z(I)$ in \hat{G} as the set of common zeros of the Fourier transforms of elements of I . We ask whether, conversely, I is characterized by $Z(I)$. This question is the problem of **spectral synthesis**, and many important results have been obtained. The statement that I must coincide with R when $Z(I)$ is empty is a formulation of the **generalized Tauberian theorem** of N. Wiener. A considerable simplification of the proof was accomplished by using the theory of Banach algebras, which was the first application of the theory (I. M. Gel'fand).

M. Holomorphic Functional Calculus [16,171]

Let R be a unital commutative Banach algebra with maximal ideal space $\mathfrak{M} = \mathfrak{M}(R)$. We define the **joint spectrum** $\text{Sp}(X)$ of a finite n -tuple $X = \{x_1, \dots, x_n\}$ in R by $\{(\hat{x}_1(\varphi), \dots, \hat{x}_n(\varphi)) \mid \varphi \in \mathfrak{M}\}$, a nonempty compact subset of \mathbb{C}^n . Then the **holomorphic functional calculus** says the following: For each $X = \{x_1, \dots, x_n\} \subset R$ there exists a unique algebra homomorphism Φ_X of the algebra $H(\text{Sp}(X))$ of \dagger germs of holomorphic functions on $\text{Sp}(X)$ into R with the following properties: (1) $\Phi_X(1) = 1$, the unity element of R , and $\Phi_X(z_i) = x_i$, $1 \leq i \leq n$; (2) if $n < m$, $X' = \{x_1, \dots, x_n\}$ and $\tilde{F}(z_1, \dots, z_m) = F(z_1, \dots, z_n)$ with $F \in H(\text{Sp}(X))$, then $\Phi_{X'}(\tilde{F}) = \Phi_X(F)$; (3) if $\{F_k \mid k = 1, 2, \dots\}$ are holomorphic in a fixed neighborhood U of $\text{Sp}(X)$ and $F_k \rightarrow F$ uniformly on U , then $\Phi_X(F_k) \rightarrow \Phi_X(F)$ (L. Waelbroeck). From this follows the **implicit function theorem**: Let $x_1, \dots, x_n \in R$, $f \in C(\mathfrak{M})$, and let $F(w, z, \dots, z_n)$ be holomorphic in a neighborhood of the set $\text{Sp}(f, x_1, \dots, x_n) = \{(f(\varphi), \hat{x}_1(\varphi), \dots, \hat{x}_n(\varphi)) \mid \varphi \in \mathfrak{M}\}$. Suppose that

$F(f, \hat{x}_1, \dots, \hat{x}_n) = 0$ but $\partial F / \partial w \neq 0$ on $\text{Sp}(f, x_1, \dots, x_n)$. Then there exists a unique $y \in R$ such that $\hat{y} = f$ and $F(y, x_1, \dots, x_n) = 0$ (R. Arens and A. Calderon). As an application we obtain Shilov's **idempotent theorem**: If K is a closed-open subset of \mathfrak{M} , there exists $x \in R$ with $\hat{x} \equiv 1$ on K and $\hat{x} \equiv 0$ on $\mathfrak{M} - K$. Set $Q(R) = \{x \in R \mid e^{2\pi i x} = 1\}$. Then $Q(R)$ is an additive subgroup of R and Shilov's theorem says that the Gel'fand representation gives an isomorphism of $Q(R)$ onto $Q(C(W)) (\cong H^0(\mathfrak{M}, \mathbb{Z}))$. Another theorem of this sort is the **Arens-Royden theorem**: $R^{-1} / \exp(R) \cong C(\mathfrak{M})^{-1} / \exp(C(\mathfrak{M})) \cong H^1(\mathfrak{M}, \mathbb{Z})$, where R^{-1} is the set of invertible elements of R and $\exp(R) = \{e^x \mid x \in R\}$. Further extensions related to K-theory have been given by J. L. Taylor and others. (For functional calculus for one variable \rightarrow 25 1 Linear Operators)

References

- [1] C. E. Rickart, General theory of Banach algebras, Van Nostrand, 1960.
- [2] M. A. Naimark, Normed algebras, Wolters-Noordhoff, 1972. (Original in Russian, 1968.)
- [3] L. H. Loomis, An introduction to abstract harmonic analysis, Van Nostrand, 1953.
- [4] K. Yosida, Functional analysis, Springer, third edition, 1971.
- [5] W. Arveson, An invitation to C*-algebra, Springer, 1976.
- [6] J. Dixmier, Les C*-algèbres et leurs représentations, Gauthier-Villars, 1964.
- [7] G. K. Pedersen, C*-algebras and their automorphism groups, Academic Press, 1979.
- [S] S. Sakai, C*-algebras and W*-algebras, Springer, 1971.
- [9] M. Takesaki, Theory of operator algebras 1, Springer, 1979.
- [10] O. Bratteli and D. W. Robinson, Operator algebras and quantum statistical mechanics, Springer, 1979.
- [11] G. A. Elliott, On the classification of inductive limits of sequences of semisimple finite-dimensional algebras, J. Algebra, 38 (1976), 29-44.
- [12] R. G. Douglas, Extensions of C*-algebras and algebraic topology, Proc. Internat. Congress Math. Helsinki 1978, vol. 2, 1980, 6999-704.
- [13] C. A. Akemann and G. K. Pedersen, Central sequences and inner derivations of separable C*-algebras, Amer. J. Math., 101 (1979), 1047-1061.
- [14] S. Sakai, Recent developments in the theory of unbounded derivations in C*-

algebras, Proc. Internat. Congress Math. Helsinki 1978, vol. 2, 1980, 709–714.

[15] W. Rudin, Fourier analysis on groups, Interscience, 1962.

[16] T. W. Gamelin, Uniform algebras, Prentice-Hall, 1969.

[17] J. L. Taylor, Banach algebras and topology, Algebra in Analysis, J. H. Williamson (ed.), Academic Press, 1975, 118186.

37 (X11.3) Banach Spaces

A. General Remarks

The notion of **Banach space** was introduced in 1922 by S. Banach and N. Wiener, working independently. The idea was to apply topological and algebraic methods to fundamental problems of analysis, such as mapping problems in infinite-dimensional function spaces (- 168 Function Spaces, 197 Hilbert Spaces, 251 Linear Operators).

B. Definition of Banach Spaces

We associate to each element x of a \dagger linear space X over the real (or complex) number field a real number $\|x\|$ satisfying the following conditions: (i) $\|x\| \geq 0$ for all x , and $\|x\| = 0$ is equivalent to $x = 0$; (ii) $\|\alpha x\| = |\alpha| \cdot \|x\|$ for any real (complex) number α ; (iii) $\|x + y\| \leq \|x\| + \|y\|$. Then $\|x\|$ is called the **norm** of the vector x , and X is called a **normed linear space**. The norm is thus an extension of the notion of the length of a vector in a \dagger Euclidean space. A normed linear space X is a \dagger metric space under the distance $\rho(x, y) = \|x - y\|$. We write $\lim_{n \rightarrow \infty} x_n = x$ or simply $x_n \rightarrow x$ when $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$ and say that x_n **converges strongly** to x . If this metric space X is \dagger complete, then X is called a **Banach space**. In normed linear spaces the addition and the multiplication by scalars are continuous. A closed linear subspace M of a Banach space X is again a Banach space, and the \dagger quotient space X/M becomes a Banach space if the norm of a \dagger coset is defined by $\|x + M\| = \inf_{m \in M} \|x + m\|$. The subset $\{x \mid \|x\| \leq 1\}$ of a normed linear space X is called the (closed) **unit ball** (or unit sphere) of X .

Examples. \dagger Function spaces C, L_p ($1 \leq p < \infty$), M, W_p^l, H_0^l , \dagger sequence spaces c, l, m and BV are all Banach spaces (- 168 Function Spaces).

C. Linear Operators and Linear Functionals

Suppose that a linear subspace $D(T)$ of a linear space X is the (definition) **domain** of a mapping T with values in a linear space X_1 . T is called a **linear operator** if $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$ for any scalars α, β and $x, y \in D(T)$. $R(T) = \{Tx \in X_1 \mid x \in D(T)\}$ is called the **range** of T . In the special case where X_1 is the real or complex number field, T is called a **linear functional**. If X and X_1 are both normed linear spaces, then T is **continuous** if and only if $s\text{-}\lim_{n \rightarrow \infty} Tx_n = Tx$ whenever $s\text{-}\lim_{n \rightarrow \infty} x_n = x$. This is equivalent to the condition $\sup_{x \in D(T), \|x\| \leq 1} \|Tx\| < \infty$. In particular, if $D(T) = X$, the linear operator T is **continuous** if and only if the set $\{Tx \mid \|x\| \leq 1\}$ is bounded. In this case, T is called a **bounded linear operator**, and $\|T\| = \sup_{\|x\| \leq 1} \|Tx\|$ is called the **norm** of the operator T . In particular, a linear operator T satisfying $\|T\| \leq 1$ is called a **contraction**. (Sometimes T is called a contraction only when $\|T\| < 1$. In that case, an operator with $\|T\| \leq 1$ is called a **nonexpansive operator**.) The **scalar multiple**, **sum**, and **product** of linear operators are defined by $(\alpha T)x = \alpha(Tx)$, $(T + S)x = Tx + Sx$, and $(ST)x = S(Tx)$, respectively. The **identity operator** I in X is defined by $I \cdot x = x$ for all $x \in X$. If the inverse mapping T^{-1} of $x \rightarrow Tx$ exists, then it is called the **inverse operator** of T . A normed linear space X is said to be **isomorphic** to a normed linear space Y if there exists a bounded linear operator T from X onto Y with bounded inverse. If T can be chosen to be isometric (i.e., $\|Tx\| = \|x\|$ for all $x \in X$), then X is said to be **isometrically isomorphic** to Y .

D. The Dual Space and the Dual Operator

The totality of continuous linear functionals f defined on a normed linear space X is a Banach space X' under the previously defined linear operations and the norm $\|f\| = \sup_{\|x\| \leq 1} |f(x)|$. This X' is called the **dual** (or **conjugate**) space of X . In view of useful properties of the \dagger inner product in \dagger Hilbert spaces, it is sometimes convenient to write $\langle x, f \rangle$ for $f(x)$. Let X and Y be normed linear spaces, and let T be a linear operator with the dense domain $D(T)$ in X and the range $R(T)$ in Y . If (f, g) is a pair with $f \in Y'$ and $g \in X'$ satisfying the equation $(Tx, f) = \langle x, g \rangle$ for any $x \in D(T)$, then g is determined uniquely by f . The operator T' defined by $T'f = g$ is linear and is called the **dual operator** (or **conjugate** or **adjoint operator**) of T . This is an extension of the notion of the \dagger transpose of a matrix in matrix theory. If T is a bounded linear operator then

T' is also a bounded linear operator such that $\|T'\| = \|T\|$.

E. The Weak Topology and the Strong Topology

Let X be a normed linear space and X' its dual space. Take a finite number of elements x'_1, x'_2, \dots, x'_n from X' , and consider the subset of $X: \{x \in X \mid \sup_{1 \leq i \leq n} |\langle x, x'_i \rangle| \leq \varepsilon\}$, $\varepsilon > 0$. If we take the totality of such subsets of X as a fundamental system of neighborhoods of 0 of X , then X is a locally convex topological linear space, denoted sometimes by X_w . This topology is called the **weak topology** of X . If a sequence $\{x_n\} \subset X$ converges to $x \in X$ with respect to the weak topology of X , then it is said to **converge weakly**. This is equivalent to the convergence $(x_n, f) \rightarrow (x, f)$ for any $f \in X'$. The original topology of X determined by the norm is then called the **strong topology** of X , and to stress the strong topology we sometimes write X_s in place of the original X . Take a finite number of elements x_1, x_2, \dots, x_n from X , and consider the subset of $X': \{x' \in X' \mid \sup_{1 \leq i \leq n} |\langle x_i, x' \rangle| \leq \varepsilon\}$, $\varepsilon > 0$. If we take the totality of such subsets of X' as a fundamental system of neighborhoods of 0 of X' , then X' is a locally convex topological linear space. We write this space as X'_{w*} and call the topology the **weak* topology** of X' . The topology of X' defined by the norm $\|f\|$ is called the **strong topology** of X' , and to stress the strong topology we write X'_s . The terms "weak" and "weakly" are used in reference to the weak topology, for instance, weak closedness and weakly closed. Similar conventions are used for the weak* and strong topologies.

The unit ball of the dual space of a normed linear space X is weak*-compact (**Banach-Alaoglu theorem**). Then by the Krein-Milman theorem (\rightarrow 424 Topological Linear Spaces) the unit ball of X' is the weak* closure of the convex hull of its extreme points. If X is a Banach space, a convex subset K of X' is weak*-closed if and only if the intersection of K and each weak*-compact subset is weak*-closed (**Krein-Shmul'yan theorem**).

F. The Hahn-Banach Extension Theorem

Let M be a linear subspace of a real linear space X and $p(x)$ a real-valued functional defined on X such that $p(x+y) \leq p(x) + p(y)$ and $p(\lambda x) = \lambda p(x)$ for all $x, y \in X$ and $\lambda \geq 0$. If a linear functional f_1 defined on M satisfies $f_1(x) < p(x)$ for all $x \in M$, then there is a linear functional f on X which extends f_1 and satisfies $f(x) \leq p(x)$ for all $x \in X$ (the **Hahn-Banach**

(extension) theorem). The Hahn-Banach extension theorem has many applications. (1) Let M be a linear subspace of a normed linear space X . Then for any $f_1 \in M'$, we can construct an $f \in X'$ such that $f(x) = f_1(x)$ for all $x \in M$ and $\|f\| = \|f_1\|$. (2) For any $x_0 \neq 0$ of a normed linear space X , we can construct an $f_0 \in X'$ such that $f_0(x_0) = \|x_0\|$ and $\|f_0\| = 1$. (3) For any closed linear subspace M of a normed linear space X and a point $x_0 \notin M$ we can construct $f_0 \in X'$ such that $\|f_0\| = 1$, $f_0(x_0) = \inf_{m \in M} \|x_0 - m\|$ and $f_0(x) = 0$ for all $x \in M$. A proposition more general than (3) is **Mazur's theorem**, which is useful in applications: (4) Let a closed subset M of a normed linear space X be convex. Then for any $x_0 \notin M$ we can construct an $f_0 \in X'$ such that $f_0(x_0) > f_0(x)$ for all $x \in M$. By (4) we can prove, e.g., that a convex set of a normed linear space is weakly closed if it is strongly closed. This proposition has the following corollaries: (5) A convex set A containing 0 is closed if and only if $(A^\circ)^\circ = A$, where A° is the polar of A (the **bipolar theorem**). (6) For two closed convex sets A and B , containing 0, the polar of $A \cap B$ coincides with the weak* closure of the convex hull of $A^\circ \cup B^\circ$. For complex linear spaces, most of the propositions in this section are valid with $\operatorname{Re} f(x)$ instead of $f(x)$.

The Hahn-Banach theorem can also be employed to prove the existence of the **generalized limit** (or **Banach limit**) $\lim_{n \rightarrow \infty} \zeta_n$ defined for all bounded real sequences $\{\zeta_n\}$ such that $\liminf_{n \rightarrow \infty} \zeta_n \leq \lim_{n \rightarrow \infty} \zeta_n \leq \limsup_{n \rightarrow \infty} \zeta_n$ and $\lim_{n \rightarrow \infty} (\alpha \zeta_n + \beta \eta_n) = \alpha \lim_{n \rightarrow \infty} \zeta_n + \beta \lim_{n \rightarrow \infty} \eta_n$.

G. Duality in Normed Linear Spaces

An element x_0 of a normed linear space X gives rise to an element x'_0 of (X'_s) determined by $\langle x_0, x' \rangle = \langle x'_0, x'_0 \rangle$ for all $x' \in X'_s$. If we write $x'_0 = Jx_0$, then J is a linear operator satisfying $\|Jx_0\| = \|x_0\|$ by (2), and so the space X is isometrically isomorphic to a linear subspace of $(X'_s)'_s$. If X_s coincides with $(X'_s)'_s$ under this isomorphism, we call X a **reflexive** (or **regular**) **Banach space**. A necessary and sufficient condition for the normed linear space X to be reflexive is that the unit ball of X be weakly compact. A convenient criterion for the reflexivity of X is that any bounded sequence $\{x_n\}$ of X contains a subsequence weakly convergent to a point of X (**Eberlein-Shmul'yan theorem**). In this connection, a Banach space X is reflexive if and only if each $x' \in X'$ attains its norm, i.e., there is an $x_0 \in X$ such that $\|x_0\| = 1$ and $|\langle x_0, x' \rangle| = \|x'\|$ (**James's theorem**).

A normed linear space is said to be **uniformly convex** if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|x\| \leq 1$, $\|y\| \leq 1$ and $\|x - y\| \geq \varepsilon$

implies $\|x + y\| \leq 2 + \delta$. A normed linear space is said to be **uniformly smooth** if for any $\varepsilon > 0$ there is a $\delta > 0$ such that $\|x\| = 1$ and $\|y\| \leq \delta$ imply $\|x + y\| + \|x - y\| \leq 2 + \varepsilon\|y\|$. A Banach space X is uniformly convex (resp. uniformly smooth) if and only if the dual X' is uniformly smooth (resp. uniformly convex). If a Banach space X is uniformly convex or smooth, it is isomorphic to a Banach space that is uniformly convex and uniformly smooth (P. Enflo). The space L_p with $1 < p < \infty$ is uniformly convex and uniformly smooth. Any uniformly convex or uniformly smooth Banach space is reflexive (**Milman's theorem**).

H. The Resonance Theorem

Let $\{T_n\}$ be a sequence of bounded linear operators from a Banach space X into a normed linear space Y . The **uniform boundedness theorem** (**resonance theorem** or **Banach-Steinhaus theorem**) states that $\sup_{n \geq 1} \|T_n\| < \infty$ if $\sup_{n \geq 1} \|T_n x\| < \infty$ for every $x \in X$. As a corollary, we have $\sup_{n \geq 1} \|x_n\| < \infty$ for any weakly convergent sequence of X . Another corollary states that the set $\{x \in X \mid \limsup_{n \rightarrow \infty} \|T_n x\| < \infty\}$ either coincides with X or is a subset of X of the † first category. This implies the so-called **principle of condensation of singularities**, which gives a general existence theorem for functions exhibiting various kinds of singularities, for example, a continuous function whose † Fourier expansion diverges at every point of a † perfect set of points having the cardinal number of the continuum.

I. The Closed Graph Theorem

Let T be a bounded linear operator from a Banach space X into a normed linear space Y . If the image of the unit ball of X under T is dense in the unit ball of Y , then for any $\varepsilon > 0$ and any $y_0 \in Y$ the equation $Tx = y$, has a solution x with $\|x\| \leq (1 + \varepsilon)\|y\|$, (**Banach's theorem**). By using the † Baire category theorem, we can derive from this the **open mapping theorem**: Every bounded linear operator from a Banach space X onto a Banach space Y maps each open set in X onto an open set in Y . As an application of the open mapping theorem, we can prove the **closed graph theorem**: A linear operator T from the whole of a Banach space into a Banach space is continuous if and only if T is a closed operator, i.e., $s\text{-}\lim_{n \rightarrow \infty} x_n = x$ and $s\text{-}\lim_{n \rightarrow \infty} Tx_n = y$ imply $Tx = y$. This theorem plays an important role in modern treatments of linear partial differential equations.

J. The Closed Range Theorem

Let X and Y be Banach spaces and T a linear closed operator with domain $D(T)$ dense in X and with range $R(T)$ in Y . Under these conditions, the following four propositions are mutually equivalent. (1) $R(T)$ is a closed set in Y . (2) $R(T')$ is a closed set in X' . (3) $R(T) = \{y \in Y \mid \langle y, y^* \rangle = 0 \text{ for all } y^* \in D(T') \text{ such that } T'y^* = 0\}$. (4) $R(T') = \{x^* \in X' \mid \langle x, x^* \rangle = 0 \text{ for all } x \in D(T) \text{ such that } Tx = 0\}$. These four propositions, as a whole, are called the **closed range theorem**. This theorem implies (5) $R(T) = Y$ if and only if T' has a continuous inverse; and (6) $R(T') = X'$ if and only if T has a continuous inverse. The following theorem is of similar nature: the following three propositions on two closed linear subspaces M and N of a Banach space are mutually equivalent. (7) $M + N$ is closed. (8) $M^{\circ} + N^{\circ}$ is strongly closed. (9) $M^{\circ} + N^{\circ}$ is weak*-closed.

The Hahn-Banach theorem, the resonance theorem, the open mapping theorem, the closed graph theorem, and the closed range theorem can be extended to various classes of † locally convex topological linear spaces. By virtue of this extension, we are able not only to treat various fundamental problems of analysis from a unified viewpoint but also to develop the theory of functional analysis itself in a new direction (\rightarrow 424 Topological Linear Spaces; concerning linear operators on a Banach space \rightarrow 68 Compact and Nuclear Operators, 251 Linear Operators, 390 Spectral Analysis of Operators).

K. Differential and Integral Calculus of Functions with Values in Banach Spaces

Calculus involving functions from a set to a Banach space is also an effective tool in various problems. A function $x(t)$ defined on an interval $[a, b]$ with values in a Banach space X is said to be **strongly (weakly) continuous** if $x(t)$ converges strongly (weakly) to $x(t_0)$ as $t \rightarrow t_0$. For a strongly (weakly) continuous function $x(t)$, the **Riemann integral** can be defined in a standard way, using strong (weak) convergence of the Riemann sum

$$\int_a^b x(t) dt = \lim \sum x(t_i')(t_{i+1} - t_i).$$

The † fundamental theorem of calculus, i.e., strong (weak) differentiability of the indefinite integral, remains true. Various definitions of integrals of a Banach space-valued function on a † measure space are discussed elsewhere (\rightarrow 443 Vector-Valued Integrals).

Now let $x(\lambda)$ be defined on a domain Ω in the complex plane with values in a complex

Banach space X . $x(\lambda)$ is said to be **holomorphic** if $f(x(\lambda))$ is holomorphic in Ω for every $f \in X'$. If $x(\lambda)$ is holomorphic, then there exists an X -valued function $y(\lambda)$ on Ω such that

$$\|\zeta^{-1}(x(\lambda + \zeta) - x(\lambda)) - y(\lambda)\| \rightarrow 0 \quad \text{as} \quad \zeta \rightarrow 0.$$

Therefore there is no difference between "strong" and "weak" in analyticity. \dagger Cauchy's integral theorem remains true for a holomorphic function $x(\lambda)$ with values in X , and the \dagger Laurent expansion

$$x(\lambda) = \sum_{n=-\infty}^{\infty} a_n(\lambda - \lambda_0)^n,$$

$$a_n = \frac{1}{2\pi i} \int_C x(\lambda)(\lambda - \lambda_0)^{-n-1} d\lambda,$$

is valid with the integral taken in the Riemannian sense. Banach space-valued holomorphic functions on complex (or real) \dagger analytic manifolds of higher dimension can be defined in a natural way by means of power series expansion. A function $\Theta(\lambda)$ defined on a domain in the complex plane with values in the Banach space $B(X, Y)$ of bounded linear operators from X to Y becomes holomorphic if $f(\Theta(\lambda)x)$ is holomorphic for every $x \in X$ and $f \in Y'$. An operator-valued holomorphic function is often called an **analytic operator function**.

L. The Approximation Property

A Banach space X is said to have the **approximation property** if there is a family $\{S_\lambda\}$ of bounded linear operators of \dagger finite rank in X such that $\inf_\lambda \|(S_\lambda - I)T\| = 0$ for all \dagger compact linear operators T in X . There is a Banach space that fails to have the approximation property (Enflo [9]). More surprisingly, the space of all bounded linear operators on an infinite-dimensional Hilbert space fails to have the approximation property (A. Szankowski). The approximation property plays a decisive role in the theory of \dagger tensor products of Banach spaces. A Banach space X is said to have the **bounded approximation property** if a family $\{S_\lambda\}$ in the definition of the approximation property can be taken bounded, i.e., $\sup_\lambda \|S_\lambda\| < \infty$. The bounded approximation property does not follow from the approximation property. The bounded approximation property is closely related to the existence of a basis. A sequence $\{e_k\}$ in a Banach space X is called a **Schauder basis** or simply a **basis** (or **hase**) for X if to each $x \in X$ there corresponds a unique sequence of numbers $\{c_k\}$ such that $\lim_{n \rightarrow \infty} \|x - \sum_{k=1}^n c_k e_k\| = 0$. Most separable Banach spaces appearing in analysis have bases. A separable Banach space has the bounded approximation property if and only if it is isomorphic to a complemented linear

subspace (\rightarrow Section N) of a Banach space with a basis (A. Pełczyński).

M. Injective and Projective Banach Spaces

Banach spaces of the type $\dagger C(\Omega)$ with compact Ω and of the type $\dagger l_1(\Omega)$ on a set Ω play a special role in the theory of Banach space. This is already seen in the fact that every Banach space is isometrically isomorphic to a subspace of a space $C(\Omega_1)$ as well as to a quotient space of a space $l_1(0, \infty)$. A Banach space X is said to be **injective** if for any Banach space Y and its linear subspace M , each bounded linear operator T from M to X can be extended to a bounded linear operator \tilde{T} from Y to X , that is, $\tilde{T}x = Tx$ for all $x \in M$. A Banach space is **injective** if and only if it is isomorphic to a complemented linear subspace of the space $C(R)$ for a compact Hausdorff space Ω with the property that the closure of every open set is open. Such a topological space is called **extremally disconnected**. The \dagger maximal ideal space of the \dagger Banach algebra $L_\infty(\Omega)$ is extremally disconnected. Hence the Banach space $L_\infty(\Omega)$ is injective. Whether every injective Banach space is isomorphic to a space $C(0)$ is still an open problem. However, a Banach space is isometrically isomorphic to a space $C(Q)$ with Ω extremally disconnected if and only if it is injective with the property of norm-preserving extension, i.e., $\|\tilde{T}\| = \|T\|$ is always possible (Nachbin-Goodner-Kelley theorem). In this connection, the following propositions on a Banach space X are mutually equivalent. (1) X' is isometrically isomorphic to a space $\dagger l_1(\mu)$. (2) For any Banach space Y , its linear subspace M , and $\varepsilon > 0$, each compact linear operator T from M to X can be extended to a compact linear operator \tilde{T} from Y to X with $\|\tilde{T}\| \leq (1 + \varepsilon)\|T\|$ (J. Lindenstrauss). A Banach space X is said to be **projective** if for any Banach space Y and its closed linear subspace M , each bounded linear operator S from X to the quotient space Y/M is lifted to a bounded linear operator \tilde{S} from X to Y , i.e., $\varphi(\tilde{S}x) = Sx$ for all $x \in X$, where φ is the quotient mapping from Y to Y/M . Projectivity of X is characterized by its being isomorphic to the space $l_1(\Omega)$ on a set Ω . A Banach space is isometrically isomorphic to a space $l_1(R)$ if and only if it is projective with the property of norm-preserving lifting, i.e., $\|\tilde{S}\| = \|S\|$ is always possible (\rightarrow [8]).

N. Complemented Subspaces Problems

A linear subspace M of a Banach space X is **complemented**, i.e., there is a closed linear

subspace N such that $M \cap N = \{0\}$ and $M + N = X$ if and only if M is the range of a bounded projection P , i.e., $P^2 = P$ and $R(P) = M$. Each nonzero closed linear subspace of a Hilbert space is complemented, or more precisely, it is the range of a projection of norm one (- 197 Hilbert Spaces). This property distinguishes Hilbert spaces from general Banach spaces: (1) A Banach space of more than 3 dimensions is isometrically isomorphic to a Hilbert space if each nonzero closed linear subspace is the range of a projection of norm one (**Kakutani's theorem**). (2) A Banach space is isomorphic to a Hilbert space if each closed linear subspace is complemented (J. Lindenstrauss and L. Tzafriri [10]).

0. Quasi-Banach Spaces and Fréchet Spaces

Let X be a linear space over the real (or complex) number field. Suppose that a real-valued function $\|x\|$ on X satisfies (i) and (ii) of Section B and (iii') $\|x + y\| \leq k(\|x\| + \|y\|)$ with a constant $k \geq 1$ independent of x and y . Then $\|x\|$ is called the **quasinorm** of x , and X equipped with a quasinorm is called a **quasinormed linear space**. Let $0 < p \leq 1$ be the root of the equation $k = 2^{(1/p)-1}$. Then there is a distance $d(x, y) = d(x - y)$ depending only on $x - y$ such that $d(x - y) \leq \|x - y\|^p \leq 2d(x - y)$. Hence a quasinormed linear space is a metric space in which a sequence x_n converges to x if and only if $\|x_n - x\| \rightarrow 0$. If a quasinormed linear space X is complete under this metric, then X is called a **quasi-Banach space**. The p -function space L_p is a quasi-Banach space for $0 < p < 1$.

If we denote by $\|x\|$ the distance $d(x - 0)$ of a quasinormed linear space, then it satisfies (i) and (iii) of Section B and (ii') $\| -x \| = \|x\|$ and $\lim_{n \rightarrow \infty} \|\alpha_n x_n - \alpha x\| = 0$ whenever $\alpha_n \rightarrow \alpha$ and $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$. A functional $\|x\|$ satisfying (i), (ii'), and (iii) is called a **pseudonorm**. If a linear space X equipped with a pseudonorm is complete, then X is called a **Fréchet space (in the sense of Banach)**. The p -function space $S(\Omega)$ is a Fréchet space. Quasinormed linear spaces and Fréchet spaces are topological linear spaces, but they need not be p -locally convex. Hence it is possible that there is no continuous linear functional except for zero. However, the open mapping theorem and the closed graph theorem hold for linear operators from a Fréchet space into a Fréchet space.

References

- [1] S. Banach, *Theorie des operations lineaires*, Warsaw, 1932 (Chelsea, 1963).
- [2] N. Dunford and J. T. Schwartz, *Linear operators I*, Interscience, 1958.

- [3] K. Yosida, *Functional analysis*, Springer, 1978.
- [4] M. M. Day, *Normed linear spaces*, Springer, 1958.
- [5] H. E. Lacey, *The isometric theory of classical Banach spaces*, Springer, 1974.
- [6] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces I and II*, Springer, 1977, 1979.
- [7] J. Diestel, *Geometry of Banach space*, Springer, 1975.
- [8] A. Pełczyński, *Some aspects of the present theory of Banach spaces*, *Oeuvres de Stefan Banach II*, Warsaw, 1979.
- [9] P. Enflo, A counterexample to the approximation problem in Banach spaces, *Acta Math.*, 130 (1973), 309-317.
- [10] J. Lindenstrauss and L. Tzafriri, On complemented subspaces problem, *Israel J. Math.*, 9 (1971), 263-269.

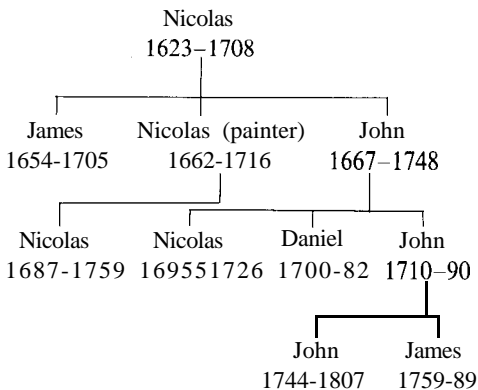
38 (XXI.1 4) Bernoulli Family

The Bernoullis, Protestants who came originally from Holland and settled in Switzerland, were a significant family to the mathematics of the 17th Century. In a single Century, the family produced eight brilliant mathematicians, all of whom played important roles in the development of calculus.

The brothers James (1654-1705) and John (1667-1748) and Daniel (1700-1782), John's second son, were especially outstanding. James and John were close friends of G. W. \dagger Leibniz, with whom they exchanged the correspondence through which it might be said that calculus developed. James studied problems related to the \dagger tautochrone and \dagger brachistochrone, as well as problems in geometry, dynamics, and other fields, including the \dagger isoperimetric problem. He was the first to change the name *calculus summatoris* to *calculus integralis* (1690). His *Ars conjectandi* was published after his death in 1713; in it is found the \dagger law of large numbers, which made his name prominent in the theory of p -probability. James had little guidance, learning mathematics on his own. He was a professor of experimental physics at the University of Basel and later became a professor of mathematics. He taught mathematics to his brother John, who succeeded him as professor at the University. John's many achievements appeared in such publications of the time as *Acta eruditorum* and *Journal des savants*. In 1701, the beginnings of the \dagger calculus of variations were seen in his solution to the isoperimetric problem. He

was the first to use the term *functio*, the root of the present term *function* (1714).

Despite discord between the brothers and also between fathers and sons, the Bernoullis were ardent teachers and brilliant researchers, who instructed not only their sons but also such mathematicians as †Euler. Their achievements were numerous in consolidating the content and form of calculus and also in expanding its application. Daniel was especially outstanding in the theory of probability; he also made contributions to the field of †hydrodynamics and to the †kinetic theory of gases. The eldest John's eldest son Nicolas (1695–1726) achieved distinction as a professor of mathematics in St. Petersburg. Daniel's youngest brother, John (1710–1790), succeeded his father, John Sr., as a professor at the University of Basel. The son of John, Jr., also named John (1744–1807), was the chairman of mathematics at the Academy of Berlin. His brother, another James (1759–1789), was a professor of experimental physics at the University of Basel. Nicolas (1687–1759), a grandson of the founder Nicolas (1623–1708) and son of Nicolas the painter (1662–1716), held Galileo's old chair of mathematics at Padua from 1716 to 1719.



In this article, the first names have been given in English. The German names corresponding to James, John, and Nicolas are Jakob, Johann, and Nikolaus, respectively.

References

- [1] M. B. Cantor, *Vorlesungen über Geschichte der Mathematik III*, Teubner, 1898.
- [2] Jacobi Bernoulli, *Opera* I, II, Cramer, 1744.
- [3] Jakob Bernoulli, *Die Werke von Jakob Bernoulli*, Birkhäuser, 1969.
- [4] Jacobi Bernoulli, *Ars conjectandi*, Bâle, 1713.
- [5] Johannis Bernoulli, *Opera omnia I-IV*, M. M. Bousquet, 1742.

39 (XIV.8) Bessel Functions

A. General Remarks

Bessel functions were first introduced in order to solve †Kepler's equation concerning planetary motions and were systematically investigated by F. W. Bessel in 1824. Since then they have appeared in various problems and have become important.

B. Bessel Functions

Separating variables for the Helmholtz equation $\Delta\Psi + k^2\Psi = 0$ in terms of cylindrical coordinates, we obtain Bessel's differential equation

$$\frac{d^2 w}{dz^2} + \frac{1}{z} \frac{dw}{dz} + \left(1 - \frac{v^2}{z^2}\right)w = 0 \quad (1)$$

for the component of the radius vector. The following two linearly independent solutions of (1):

$$H_v^{(1)}(z) = \frac{1}{\pi} \int_{L_1} e^{-iz \sin \zeta + iv \zeta} d\zeta, \quad H_v^{(2)}(z) = \frac{1}{\pi} \int_{L_2} e^{-iz \sin \zeta + iv \zeta} d\zeta, \quad (2)$$

are called the **Hankel functions of the first and second kind**, respectively, where the contour L_1 of the first integration is a curve from $(-\pi + 0) + i\infty$ to $-0 - i\infty$, and L_2 is a curve from $+0 - i\infty$ to $(\pi - 0) + i\infty$. If both z and v are real, we have

$$H_v^{(1)}(\bar{z}) = H_v^{(2)}(z), \quad \overline{H_v^{(2)}(z)} = H_v^{(1)}(z), \quad (3)$$

where \bar{z} is the complex conjugate of z . Hence

$$J_v(z) = (H_v^{(1)}(z) + H_v^{(2)}(z))/2, \quad N_v(z) \equiv Y_v(z) = (H_v^{(1)}(z) - H_v^{(2)}(z))/2i \quad (4)$$

are real functions. If both z and v are complex, the functions $J_v(z)$ and $N_v(z)$ defined in (4) are also called **Bessel functions** and **Neumann functions**, respectively. The other names for $J_v(z)$, $N_v(z)$, and $H_v(z)$ are **Bessel functions of the first, second, and third kind**, respectively. Each of them satisfies the following recurrence formulas:

$$2 \frac{dC_v(z)}{dz} = C_{v-1}(z) - C_{v+1}(z), \quad (2v/z)C_v(z) = C_{v-1}(z) + C_{v+1}(z). \quad (5)$$

In general, functions satisfying the simultaneous †differential-difference equations (5) are called **cylindrical functions**. Every cylindrical

function $C_v(z)$ is represented in the form $C_v(z) = a_1(v)H_v^{(1)}(z) + a_2(v)H_v^{(2)}(z)$, where $a_1(v)$ and $a_2(v)$ are arbitrary periodic functions of period 1 with respect to v .

If $v = n$ (an integer) we have

$$\begin{aligned} J_{-n}(z) &= (-1)^n J_n(z), \\ N_{-n}(z) &= (-1)^n N_n(z), \end{aligned} \quad (6)$$

which show the linear dependency of J_{-n} and J_n , and N_{-n} and N_n , respectively. If $v \neq n$ (an integer), as the fundamental solutions of (1) we can take a pair J_v and J_{-v} , or N_v and N_{-v} . In (2) if we take as a contour of integration a curve from $(-\pi + 0) + i\infty$ to $(\pi - 0) + i\infty$, we obtain an integral representation for $J_v(z)$, which yields the relations

$$\begin{aligned} J_v(ze^{im\pi}) &= e^{imv\pi} J_v(z), \\ J_{-v}(ze^{im\pi}) &= e^{-imv\pi} J_{-v}(z). \end{aligned} \quad (7)$$

If $v = n$ (an integer) and $\operatorname{Re} z > 0$, we obtain

$$J_n(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{iz \sin \zeta + in\zeta} d\zeta \quad (8)$$

$$= \frac{1}{\pi} \int_0^{\pi} \cos(z \sin \zeta - n\zeta) d\zeta, \quad (9)$$

which is called **Bessel's integral**. These representations imply the following expansions by means of **generating functions**:

$$e^{iz \sin \zeta} = \sum_{n=-\infty}^{\infty} J_n(z) e^{in\zeta}, \quad (10)$$

$$\cos(z \sin \zeta) = J_0(z) + 2 \sum_{n=1}^{\infty} J_{2n}(z) \cos 2n\zeta, \quad (11)$$

$$\sin(z \sin \zeta) = 2 \sum_{n=0}^{\infty} J_{2n+1}(z) \sin(2n+1)\zeta.$$

Making a change of variable $u = \exp(-i\zeta)$ in (2), we obtain

$$J_v(z) = \frac{1}{2\pi i} \int_L \exp\left(\frac{z}{2}\left(u - \frac{1}{u}\right)\right) u^{-v-1} du, \quad (12)$$

where L is a contour starting at the point at infinity with the argument $-\pi$, encircling the origin in the positive direction, and tending to the point at infinity with the argument π . From (12) we obtain a power series expansion

$$J_v(z) = \frac{z^v}{2^v} \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(v+m+1)} \frac{z^{2m}}{2^m}, \quad (13)$$

obtained also from (1) by a power series expansion at $z = 0$, which is a **regular** singular point of (1). Substituting (13) into

$$N_v(z) = (\cos vJ_v(z) - J_{-v}(z))/\sin v\pi, \quad (14)$$

we obtain a power series expansion for $N_v(z)$. A power series expansion for $N_v(z)$ for an integer n is obtained by taking the limit $v \rightarrow n$ (\rightarrow Appendix A, Table 19). In particular, if $v =$

$n + 1/2$ (n is an integer), we have

$$J_{n+1/2}(z) = (-1)^n \frac{(2z)^{n+1/2}}{\sqrt{\pi}} \frac{d^n}{d(z^2)^n} \left(\frac{\sin z}{z} \right),$$

$$n = 0, 1, 2, \dots,$$

which is represented by elementary functions and is sometimes called simply the **half Bessel function**. Bessel functions for half-integers have appeared also as radius vector components when the variables in the Helmholtz equation are separated by spherical coordinates. The function

$$j_n(z) = \sqrt{\pi/2z} J_{n+1/2}(z)$$

is called the **spherical Bessel function**.

We have the following **addition theorem**:

$$H_n^{(\mu)}(k\rho) e^{in\psi} = \sum_{m=-\infty}^{\infty} J_n(kr_2) H_{n+m}^{(\mu)}(kr_1) e^{im\varphi},$$

$$\mu = 1, 2,$$

where

$$\rho = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \varphi},$$

$$\rho \cos \psi = r_1 - r_2 \cos \varphi, \quad \rho \sin \psi = r_2 \sin \varphi.$$

C. Zero Points of the Function $J_v(z)$

From the differential equations satisfied by $J_v(\alpha z)$, we have

$$\begin{aligned} (\alpha^2 - \beta^2) \int_0^1 z J_v(\alpha z) J_v(\beta z) dz \\ = \beta J_v(\alpha) J'_v(\beta) - \alpha J'_v(\alpha) J_v(\beta). \end{aligned} \quad (15)$$

By letting $\beta \rightarrow \alpha$ in (15), we have

$$\begin{aligned} \int_0^1 z (J_v(\alpha z))^2 dz \\ = \frac{1}{2} \left(\left(1 - \frac{v^2}{\alpha^2} \right) (J_v(\alpha))^2 + (J'_v(\alpha))^2 \right) \end{aligned} \quad (16)$$

If α and β are distinct roots of $J_v(z) = 0$, we have from (15)

$$\int_0^1 z J_v(\alpha z) J_v(\beta z) dz = 0, \quad \operatorname{Re} v > -1. \quad (17)$$

The integral formulas (15), (16), and (17) are called **Lommel's integrals**.

As for the zero points of $J_v(z)$, the following facts are well known: $J_v(0) = 0$ if $v > 0$. $J_v(z)$ has no multiple zero points other than $z = 0$. $J_v(-\alpha) = 0$ if $J_v(\alpha) = 0$. Every zero point of $J_v(z)$ is real if $v > -1$. Between two adjacent zero points that are positive, there exists one and only one zero point of $J_{v-1}(z)$ and $J_{v+1}(z)$, respectively. $J_v(z)$ has a countably infinite set of zero points on the real axis. When $v \geq 0$ is rational, every zero point of $J_v(z)$ except $z = 0$ is a **transcendental** number. The transcendence of π is a special case of this result for $v = 1/2$.

D. Expansion by means of Bessel Functions

Let $f(r, \varphi)$ be defined for $0 < r < 1$, $-\pi < \varphi < \pi$, and $\alpha_{n,1}, \alpha_{n,2}, \dots, \alpha_{n,s}, \dots$ ($0 < \alpha_{n,s} < \alpha_{n,s+1}$, $s = 1, 2$, for every n) be zero points of $J_n(x)$ ($n = 0, 1, 2, \dots$). Then we have an expansion

$$f(r, \varphi) = \sum_{n=0}^{\infty} \sum_{s=1}^{\infty} (a_{n,s} \cos n\varphi + b_{n,s} \sin n\varphi) J_n(\alpha_{n,s} r), \quad (18)$$

which is called the **Fourier-Bessel series**. The coefficients $a_{n,s}$ and $b_{n,s}$ are determined by the properties of \dagger Fourier series and (16) and (17) as follows:

$$\left. \begin{aligned} a_{n,s} \\ b_{n,s} \end{aligned} \right\} = \frac{\varepsilon_n}{\pi(J_{n+1}(\alpha_{n,s}))^2} \times \int_0^1 \int_{-\pi}^{\pi} f(r, \varphi) J_n(\alpha_{n,s} r) \frac{\cos n\varphi}{\sin n\varphi} r d\varphi dr,$$

$$\varepsilon_0 = 1, \quad \varepsilon_1 = \varepsilon_2 = \dots = 2.$$

The integral transformation

$$g(y) = \int_0^{\infty} x f(x) J_n(xy) dx \quad (19)$$

is called the **Fourier-Bessel transform**. If $f(x)$ is sufficiently smooth and tends rapidly to zero as $x \rightarrow \infty$, the following inversion formula holds:

$$f(x) = \int_0^{\infty} y g(y) J_n(xy) dy. \quad (20)$$

There are other types of series expansions in terms of Bessel functions as follows: **Dini's series**

$$\sum_{m=1}^{\infty} a_m J_v(\lambda_m x)$$

(λ_m is the m th positive root of $xJ'_v(x) + HJ_v(x) = 0$, where H is a real constant); **Kapteyn's series**

$$\sum_{m=1}^{\infty} a_m J_{v+m}((v+m)x);$$

Schlömilch's series

$$\frac{a_0}{2} + \sum_{m=0}^{\infty} a_m J_0(mx);$$

and the **generalized Schlömilch series**

$$\frac{1}{2} \frac{a_0}{\Gamma(v+1)} + \sum_{m=1}^{\infty} \frac{a_m J_v(mx) + b_m H_v(mx)}{(mx/2)^v},$$

where $H_v(mx)$ is the Struve function (-Section F).

E. Asymptotic Expansion

If $|z|$ or v is sufficiently large, the asymptotic representation for Bessel functions is obtained

by applying the \dagger method of the steepest descent for (2). If $|z| > |v|$, we have

$$H_v^{(1)}(z) \sim \sqrt{\frac{2}{\pi z}} \exp i \left(z - \frac{\pi}{2} v - \frac{\pi}{4} \right),$$

$$-\pi < \arg z < 2\pi,$$

$$H_v^{(2)}(z) \sim \sqrt{\frac{2}{\pi z}} \exp \left(-i \left(z - \frac{\pi}{2} v - \frac{\pi}{4} \right) \right),$$

$$-2\pi < \arg z < \pi,$$

$$J_v(z) \sim \sqrt{\frac{2}{\pi z}} \cos \left(z - \frac{\pi}{2} v - \frac{\pi}{4} \right),$$

$$-\pi < \arg z < \pi,$$

$$N_v(z) \sim \sqrt{\frac{2}{\pi z}} \sin \left(z - \frac{\pi}{2} v - \frac{\pi}{4} \right).$$

Hence $H_v^{(1)}(z)$ tends to zero as $|z| \rightarrow \infty$ in the upper half-plane, and becomes large exponentially as $|z| \rightarrow \infty$ in the lower half-plane. The results for $H_v^{(2)}$ are obtained by interchanging "upper half-plane" with "lower half-plane" in this statement.

If both $|z|$ and $|v|$ are sufficiently large, we have the **Debye asymptotic representations**. For example, if $z = v \sec \beta$ ($v > 0$, $\beta > 0$), we have

$$H_v^{(1,2)}(v \sec \beta) \sim (\pi v \tan \beta/2)^{-1/2} \times \exp(\pm i v (\tan \beta - \beta) - \pi/4).$$

If $z = v \operatorname{sech} \alpha$ ($v > 0$, $\alpha > 0$), we have

$$\begin{aligned} J_v(v \operatorname{sech} \alpha) &\sim (2\pi v \tanh \alpha)^{-1/2} \exp v (\tanh \alpha - \alpha), \\ N_v(v \operatorname{sech} \alpha) &\sim (nv \tanh \alpha/2)^{-1/2} \exp v (\alpha - \tanh \alpha), \end{aligned}$$

If $|v| \sim |z|$, we have

$$\begin{aligned} H_v^{(1,2)}(v \sec \beta) &\sim \frac{\tan \beta}{\sqrt{3}} \\ &\times \exp \left(\pm i \left(\frac{\pi}{6} + v \left(\tan \beta - \frac{1}{3} \tan^3 \beta - \beta \right) \right) \right) \\ &\times H_{1/2}^{(1,2)}((v/3) \tan^3 \beta) + O(v^{-1}), \end{aligned}$$

which is called **Watson's formula**.

F. The Wagner Function

As an application of Bessel functions to the theory of nonstationary aircraft wings, T. Theodorsen introduced the function

$$C(z) = H_1^{(2)}(z)/(H_0^{(2)}(z) + H_1^{(2)}(z))$$

[6], and H. Wagner considered the function

$$k_1(s) = \frac{1}{2\pi i} \int_{Br} e^{ws} \frac{2C(-iw)}{w} dw$$

[5], where $H_0^{(2)}(z)$, $H_1^{(2)}(z)$ are Hankel functions, and \int_{Br} means a Bromwich integral giving the inverse Laplace transform. Then $C(z)$ and $k(s)$ are called the **Theodorsen function** and **Wagner function**, respectively. The function $k(s)$ is equal to the lift coefficient when a 2-dimensional flat wing suddenly proceeds forward a distance s at an angle of incidence $1/\pi$.

G. Functions Related to Bessel Functions

The following functions are closely related to Bessel functions (- Appendix A, Table 19.IV).

(1) **Modified Bessel functions.**

$$I_\nu(z) \equiv e^{-i\nu\pi/2} J_\nu(e^{i\pi/2} z),$$

$$K_\nu(z) = \frac{\pi}{2} \frac{I_{-\nu}(z) - I_\nu(z)}{\sin \nu\pi}$$

(2) **Kelvin functions.**

$$\text{ber}_\nu(z) \pm i \text{bei}_\nu(z) \equiv J_\nu(e^{\pm 3\pi i/4} z),$$

$$\text{her}_\nu(z) \pm i \text{hei}_\nu(z) \equiv H_\nu^{(1)}(e^{\pm 3\pi i/4} z),$$

$$\text{ker}_\nu(z) \equiv -(\pi/2) \text{hei}_\nu(z),$$

$$\text{kei}_\nu(z) \equiv (\pi/2) \text{her}_\nu(z).$$

(3) **Struve function.**

$$H_\nu(z) = \sum_{m=0}^{\infty} \frac{(-1)^m (z/2)^{\nu+2m+1}}{\Gamma[m+(3/2)] \Gamma[v+m+(3/2)]}$$

(4) **Anger function.**

$$J_\nu(z) \equiv \frac{1}{\pi} \int_0^\pi \cos(\nu\theta - z \sin \theta) d\theta.$$

When ν is an integer n , we have $J_\nu(z) = J_{-n}(z)$.

(5) **H. F. Weber function.**

$$E_\nu(z) \equiv \frac{1}{\pi} \int_0^\pi \sin(\nu\theta - z \sin \theta) d\theta.$$

The last three functions satisfy certain inhomogeneous Bessel differential equations. Many other functions, such as \dagger Airy's integral, can be represented by Bessel functions.

References

[1] G. N. Watson, A treatise on the theory of Bessel functions, Cambridge Univ. Press, 1922.
[2] A. Gray and G. B. Mathews, A treatise on Bessel functions and their application to physics, Macmillan, second edition, 1922.
[3] R. Weyrich, Die Zylinderfunktionen und ihre Anwendungen, Teubner, 1937.
[4] F. Bowman, Introduction to Bessel functions, Longmans-Green, 1938.
[5] H. Wagner, Über die Entstehung des dynamische Auftriebes von Tragflügeln, Z. Angew. Math. Mech., 5 (1925), 17-35.

[6] T. Theodorsen, General theory of aerodynamic instability and the mechanism of flutter, NASA Tech. Rep. 496 (1935).

Also \rightarrow references to 167 Functions of Confluent Type, 389 Special Functions.

40 (XVIII.1 6)
Biometrics

A. General Remarks

Biometrics is the branch of science that applies mathematical and statistical methods to biological problems, and deals with all the phenomena that affect the physical, social, and psychological well-being of human beings. These phenomena involve the relationships of groups of human beings to other human beings, to animals, microbes, and plants, and to physical and chemical elements in the environment. In dealing with these problems the biometrician encounters such theoretical tasks as analyzing autocorrelated data in time series, and such practical undertakings as cost-vs-benefit evaluations of health programs.

Biometrics began in the middle of the 17th century when Sir William Petty and John Graunt developed a new method of analyzing the London Bills of Mortality. Petty and Graunt essentially invented the field of *vital statistics* by studying the reports of christenings and causes of death and proposed a method called "political arithmetic."

Some fields to which biometrics is relevant are given below.

B. Statistical Genetics

After early development in vital statistics, *statistical genetics* was founded on the new ideas emerging in statistics. Major contributions were made by Charles Darwin (1809–1882), Francis Galton (1822–1911), Karl Pearson (1857–1936), and Ronald A. Fisher (1890–1962).

Galton was the first to use the term "regression" in statistics (\rightarrow 403 Statistical Models D), when he observed that sons regressed linearly on their fathers with respect to stature. He called the phenomenon a "regression to mediocrity" because the deviations of the stature of sons were less than those of fathers. This gave rise to the measurement of correlation in the bivariate normal distribution by means of the coefficient of correlation (Pearson, 1897). Pearson is credited with the

creation of the discipline of biometry (biometrics), and he established the journal *Biometrika* to promote studies in the field. Fisher's major contributions were to genetics and statistical theories. His *General theory of natural selection* appeared in 1930. This landmark book, along with earlier and later publications, represented Fisher's attempts to give quantitative form to Darwin's views and to frame a statistical theory of evolution.

C. Bioassay

Bioassay is a set of techniques for evaluating the effectiveness of dosages of drugs by monitoring biological responses. It entails the use of special transformations, such as probits and logits, as well as the application of regression to the estimation of dosages that are p percent effective within stated confidence limits. Problems to be solved include measuring relative potency, slope-ratio assays, and quantal responses vis-à-vis tolerance distributions.

D. Demography

Demography, which includes traditional vital statistics, rates and ratios, life tables, competing risks, actuarial statistics, and census enumeration techniques, is a part of biometrics. In this category, many tabulations of data consist of time series of events or rates classified by age. For the analysis of such data, the cohort analysis techniques described in Hastings and Berry [3] are employed.

E. Epidemiology

The quantitative description of an epidemic should state the sensitivity and specificity of any diagnostic tests, as well as the true incidence or prevalence of the epidemic from survey results. Within an epidemiological theory, a disease is studied by the use of deterministic and stochastic models, wherein the theory of Markov chains can be applied (- 260 Markov Chains). Differential equations involving probability-generating functions or moment-generating functions can be solved to yield the probability distribution, \bar{x} mean, and σ^2 variance of the number of infected individuals as functions of time [4, 5]. Fundamental to this whole field of application is a clear understanding of causality and association.

When clinical trials are possible, two groups of persons, "treated" and "untreated," are monitored over a period of time with regard to the incidence or recovery from the disease

under study. The techniques in this procedure include compiling reports on persons to be observed, using double-blind techniques, and combining multiple-response variables by use of multivariate analysis (- 280 Multivariate Analysis).

If clinical intervention is forbidden on ethical grounds, e.g., studying congenital malformation by infecting pregnant women with German measles, the relative risk of exposure is estimated from retrospective studies. In practice, almost all statistical studies in epidemiology are retrospective with the sole exception of clinical trials. The research is ex post facto because investigators are mostly confined to describing and analyzing sudden and/or obvious events in the etiology of the disease.

F. Clinical Trials

In clinical trials, many problems have arisen for which biometricians have had to develop special techniques. One such technique takes account of unexpected adverse effects of drugs, and the consequent early termination of a trial. Moreover, when data demonstrate a trend earlier than expected, investigators will desire to end the accession of patients and to stop further treatment with what may be an inferior regimen. This means that the biometrician must be familiar with the problems of multiple examinations of data, multiple comparisons, and other adjustment procedures required by the ex post facto dredging of data.

G. Future Trends

There are two areas in which the biometrician has played a leading role recently. These are pertinent in many different applications and problems, and considerable methodological research has been devoted to the two areas. The areas are "mathematical modeling" and "effects of hazardous substances."

Mathematical Modeling. The relationship between a set of independent variables and the dependent or response variable(s) is usually referred to as a mathematical model. The model may take the form of a standard multiple regression analysis with a single response variable or with multiple response variables as in multivariate analysis (- 280 Multivariate Analysis).

It is generally assumed that specialists with substantive knowledge of the specific field of applications (epidemiology, toxicology, pharmacology, radiology, genetics, etc.) play a

crucial role in determining a model or relationship between a set of independent variables and response variable(s). However, it is mainly the biometrician who finally selects the specific model establishing the functional relationship and who attempts to measure the strength or influence of the independent variables therein. The biometrician is also expected to contribute substantially to the decision as to whether the relationship is causal, or merely one of association or correlation. For example, in the measurement of carcinogenicity of a food additive or drug, questions arise as to whether a substance can be judged harmful if it "accelerates" the appearance of a tumor even though it does not increase the incidence of the abnormal growth. In general, the answer to this question is in the affirmative when an unequivocal dosage-response relationship is indicated between the substance and the tumor.

Effects of Hazardous Substances. With the successful conquest of most of the infectious diseases that have plagued mankind throughout history, health authorities have recently been concentrating on two chronic diseases whose etiology is yet to be determined: cardiovascular disease and cancer. In both cases there is no disagreement with the thesis that heredity exercises a determining influence, but the role of the environment in causing many cases is also unquestioned. Measurement of the risk due to potentially toxic substances in the environment, principally with respect to these two diseases, represents the greatest challenge to the biometrician today. The social benefits of success make this a tantalizing area of research, though exceptional complexities are involved.

For related topics → 263 Mathematical Models in Biology.

References

- [1] D. J. Finney, Statistical method in biological assay, Griffin, second edition, 1964.
- [2] D. R. Cox, The analysis of binary data, Methuen, 1970.
- [3] D. W. Hastings and L. G. Berry (eds.), Cohort analysis: A collection of interdisciplinary readings, Scripps Foundation for Research in Population Problems, Oxford, Ohio, 1979.
- [4] N. T. J. Bailey, The mathematical theory of epidemics, Griffin, 1957.
- [5] M. S. Bartlett, Stochastic population models in ecology and epidemiology, Methuen, 1960.

41 (XX.20) Boltzmann Equation

A. Introduction

The Boltzmann equation is an equation of motion of a rarefied gas given by L. Boltzmann in 1872 [1]. He used it successfully in his pioneering work on the kinetic theory of gases and on the more general statistical mechanics (- 402 Statistical Mechanics), but the equation was held to be inconsistent with the †classical mechanics used in its derivation; objections were raised by, e.g., J. Loschmidt (1895) and E. Zermelo (1896). Through the controversies it became widely recognized that the Boltzmann equation should be justified by means of the †theory of probability, and many such justifications have been proposed. Also, this equation has been studied extensively as a significant nonlinear partial differential equation.

B. Boltzmann Equation

Let $f = f(t, x, \xi)$ be the density of gas molecules having position $x \in \mathbf{R}^3$ and velocity $\xi \in \mathbf{R}^3$ at time t . The Boltzmann equation is a conservation law for f of the form [1-3]

$$f_t = -\xi \cdot \nabla_x f - a(t, x) \cdot \nabla_\xi f + Q[f]. \quad (1)$$

Here $a(t, x)$ denotes the external force and Q is a quadratic nonlinear integral operator in ξ -space describing binary collisions of the molecules. The integral kernel of Q , called the collision cross section, depends on the intermolecular forces. The two classical examples are the hard ball model (a gas of rigid spheres) and the inverse power law potential (one proportional to r^{-s} , where r is the intermolecular distance and $s > 1$). The latter gives rise to singularities in the kernel, and a cutoff is customarily employed to avoid this difficulty. Grad's hard ($s \geq 5$) and soft ($s < 5$) cutoff potentials may be effectively used in the study of (1) [3].

If the gas is contained in a vessel (domain), then f must also satisfy boundary conditions determined by the assumed law of reflection at the walls (boundaries) (specular reflection, random reflection, etc.). For spatially homogeneous gases with no external forces, (1) reduces to

$$f_t = Q[f], \quad (2)$$

with $f = f(t, \xi)$. The Maxwellians, those solutions satisfying the †Maxwell-Boltzmann distribution law, are the only stationary solutions of (2); they describe equilibrium states.

C. Justification by the Theory of Probability

Consider (2). The collision process is a †Markov process, and the †master equation obtained from the †Chapman-Kolmogorov equality is equivalent to (2) under the assumption of “propagation of molecular chaos” (G. Uhlenbeck (1942)). This assumption becomes valid in the limit as N (total number of molecules) $\rightarrow \infty$ (M. Kac (1955), H. McKean (1967), F. Grünbaum (1971)) [4]. Also, non-linear Markov processes defined by (2) (with probability density f/N) have been studied (McKean (1967), H. Tanaka [5]). No such results have yet been established for (1).

D. Existence of Solutions

T. Carleman [2] gave the first solutions of the Boltzmann equation, solving globally in time the †Cauchy problem for (2) for the hard ball model. His result has been extended to a wide class of potentials.

The spatially inhomogeneous case (1) is also known to have global solutions if $a(t, x) = 0$ and if the initial data for f are nearly Maxwellian. This was first proved by S. Ukai [6] assuming periodicity in x , and then proved for the Cauchy problem by Ukai [7] and T. Nishida and K. Imai [8], and for the initial boundary problem by J. P. Guiraud [9] (random reflection) and Y. Shizuta and K. Asano [10] (specular reflection) for bounded domains, and by Ukai and Asano [11] for exterior domains, all assuming Grad's hard cutoff potentials. The case of soft cutoff potentials was also solved (Ukai and Asano, R. Caflisch (1980)). All the solutions are unique and tend to Maxwellians as $t \rightarrow \infty$, with certain decay rates. There are some results also on 1-dimensional shock wave solutions (B. Nicolaenko (1975)) and stationary solutions (J. P. Guiraud (1972) (bounded domains), Ukai and Asano (1980) (exterior domains)).

For initial data far from the Maxwellian, (1) remains unsolved even locally in time for hard cutoff potentials. As for noncutoff potentials, the existence theorem in [5] is the only result known so far for either (1) or (2).

E. Hilbert Expansion

Put $f^\varepsilon = f/N$ with $\varepsilon = 1/N$. If f^ε has a power series expansion in ε , then (1) gives an infinite system of equations for the coefficients. D. Hilbert (1912) proposed a method of solving this system as an application of his theory of integral equations, initiating the attempt to

solve the Boltzmann equation. No proofs of convergence exist, but his expansion and the improved one by S. Chapman (1917) and D. Enskog (1917) made possible the applications of (1) to hydrodynamics. The zeroth-order approximation to the expansion gives rise to the compressible Euler equation, and the first-order approximation yields the compressible Navier-Stokes equation. The solutions f^ε of the Cauchy problem for (1) with nearly Maxwellian initial data converge to those of the compressible Euler equation locally in t as $\varepsilon \rightarrow 0$ (Nishida [12], Caflisch (1980)), and to those of the compressible Navier-Stokes equation asymptotically as $t \rightarrow \infty$ when $\varepsilon > 0$ is fixed [13].

References

- [1] L. Boltzmann, Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen, Wien. Ber., 66 (1872), 275–370.
- [2] T. Carleman, Problèmes mathématiques dans la théorie cinétique des gaz, Almqvist and Wiksell, 1957.
- [3] H. Grad, Asymptotic theory of the Boltzmann equation, Rarefied Gas Dynamics, J. A. Laurmann (ed.), Academic Press, 1963, vol. 1.
- [4] M. Kac, Some probabilistic aspects of the Boltzmann equation, Acta Physica Austriaca, suppl. x (1973), 379–400.
- [5] H. Tanaka, On Markov process corresponding to Boltzmann's equation of Maxwellian gas, Japan-USSR Symp. on Probability Theory, 2nd, Kyoto, 1972, G. Maruyama and Yu. V. Prokhorov (eds.): Lecture notes in math. 330, Springer, 1973.
- [6] S. Ukai, On the existence of global solutions of mixed problem for the Boltzmann equation, Proc. Japan Acad., 50 (1974), 179–184.
- [7] S. Ukai, Les solutions globales de l'équation de Boltzmann dans l'espace tout entier, C. R. Acad. Sci. Paris, 282A (1976), 317–320.
- [8] T. Nishida and K. Imai, Global solutions to the initial value problem for the nonlinear Boltzmann equation, Publ. Res. Inst. Math. Sci., 12 (1976), 229–239.
- [9] J. P. Guiraud, An H theorem for a gas of rigid spheres in a bounded domain, Théories Cinétiques Classiques et Relativistes, Colloq. CNRS, 1975.
- [10] Y. Shizuta and K. Asano, Global solutions of the Boltzmann equation in a bounded domain, Proc. Japan Acad., 53A (1977), 3–5.
- [11] S. Ukai and K. Asano, On the initial boundary value problem of the linearized Boltzmann equation in an exterior domain, Proc. Japan Acad., 56A (1980), 12–17.

[12] T. Nishida, Fluid dynamical limit of the Boltzmann equation to the level of the compressible Euler equation, *Comm. Math. Phys.*, 61 (1978), 119-148.

[13] S. Kawashima, A. Matsumura, and T. Nishida, On the fluid dynamical approximation to the Boltzmann equation at the level of the Navier-Stokes equation, *Comm. Math. Phys.*, 70 (1979), 97-124.

42 (11.15) Boolean Algebras

A. Boolean Algebras

Boolean algebra was introduced by G. Boole to study logical operations (- 411 Symbolic Logic). It is now included within the more general concept of †lattice or lattice-ordered set (- 243 Lattices) and appears not only in logic but also very often in analysis in the form of a particular lattice of sets, e.g., the lattice of †measurable sets.

Let L be a given set and suppose that to any pair of its elements x, y there correspond two elements $x \cup y, x \cap y$ of L (called **join** and **meet** of x and y , respectively) such that the following laws are valid: (1) $x \cup y = y \cup x, x \cap y = y \cap x$ (commutative law); (2) $x \cup (y \cap z) = (x \cup y) \cap z, x \cap (y \cup z) = (x \cap y) \cup z$ (associative law); (3) $x \cup (y \cap x) = (x \cup y) \cap x = x$ (absorption law); (4) $x \cup (y \cap z) = (x \cup y) \cap (x \cup z), x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$ (distributive law); ($x, y, z \in L$). From (1), (2), and (3) it follows further that $x \cup x = x \cap x = x$ (idempotent law). If $x \leq y$ is defined to mean $x \cup y = y$, L becomes an †ordered set with respect to the ordering \leq . Now suppose, moreover, that the following law holds: (5) there exist a least element 0 and a greatest element 1 , and for any element x there exists an element x' satisfying $x \cup x' = 1, x \cap x' = 0$ (**law of complementation**). Then L is called a **Boolean algebra** (or †**Boolean lattice**). In this case x' is uniquely determined by x and is called the **complement** of x . The binary operations $(x, y) \mapsto x \cup y, x \cap y$ together with the operation $x \mapsto x'$ are called **Boolean operations**. These operations obey **de Morgan's law** $(x \cup y)' = x' \cap y', (x \cap y)' = x' \cup y'$.

B. Generalized Boolean Algebras

Suppose that $a \leq b$ holds for two given elements a, b of an ordered set. Then the set of all elements x satisfying $a \leq x \leq b$ is denoted by $[a, b]$ and is called an **interval**. An interval of a

Boolean algebra is also a Boolean algebra with respect to the induced operations \cup and \cap , where the least and greatest elements are a and b , respectively, and the complement of x in $[a, b]$ is equal to $a \cup (x' \cap b) = (a \cup x') \cap b$. More generally, if a set L with two operations \cup, \cap satisfying (1)–(4) above has a least element 0 and if each interval of L satisfies (5) (i.e., is a Boolean algebra), then L is called a **generalized Boolean algebra**.

C. Boolean Rings

A ring L satisfying the condition $xx = x$ for all $x \in L$ (i.e., all of its elements are †idempotent) is called a **generalized Boolean ring**, and if it has a unity element then it is called a **Boolean ring**. A generalized Boolean ring L satisfies $x + x = 0$ for all $x \in L$ and is necessarily a commutative ring. A (generalized) Boolean algebra L becomes a (generalized) Boolean ring if for any elements x, y of L the sum $x + y$ is defined to be the complement of $x \cap y$ in the interval $[0, x \cup y]$, and the product xy is defined to be $x \cap y$. A nonempty subset J of a (generalized) Boolean algebra L is an ideal with respect to the corresponding structure of the ring if and only if $x \cup y \in J$ for $x, y \in J$ and $x \cap y \in J$ for $x \in J, y \in L$. More generally, in any lattice, a nonempty subset that satisfies these conditions is sometimes called an **ideal** of the lattice.

D. Representation of a Boolean Algebra

Any Boolean algebra L is isomorphic to a Boolean lattice of subsets in a set X . If L is of finite height, then L is isomorphic to the Boolean lattice $\mathfrak{P}(X)$ of all subsets of X . In general X can be taken to be the set of all maximal ideals of L . Let $a \in L$ and let $O(a) = \{m \mid m \in X, a \notin m\}$. The isomorphism is obtained by the mapping $a \mapsto O(a)$. If we define a topology in X such that $\{O(a) \mid a \in L\}$ is the †open base, then X is a compact, totally disconnected T_1 space and $O(a)$ is characterized as a compact open set in X . Such a space X is called a **Boolean space** (M. H. Stone [3, 4]).

In any complete Boolean algebra L , the complete distributive laws hold: $(\sup_I x_i) \cap y = \sup_I (x_i \cap y)$ and its dual. These are equivalent to the stronger relations: $(\sup_I x_i) \cap (\sup_J y_j) = \sup_{IJ} (x_i \cap y_j)$ and its dual. In order for a Boolean algebra L to be isomorphic to the Boolean algebra $\mathfrak{P}(X)$ of all subsets of X , it is necessary and sufficient that the following strongest complete distributive laws hold: $\inf_I (\sup_{J(i)} x_{ij}) = \sup_F (\inf_{I, \varphi(i)} x_{i, \varphi(i)})$ (where F is the set of all functions φ assigning to each $i \in I$ a value $\varphi(i) \in J(i)$) and its dual.

References

- [1] P. R. Halmos, Lectures on Boolean algebras, Van Nostrand, 1963.
- [2] R. Sikorski, Boolean algebras, Erg. Math., Springer, second edition, 1964.
- [3] M. H. Stone, The theory of representations for Boolean algebras, Trans. Amer. Math. Soc., 40 (1936), 37–111.
- [4] M. H. Stone, Applications of the theory of Boolean rings to general topology, Trans. Amer. Math. Soc., 41 (1937), 375–481.
- [5] G. Boole, Collected logical works 1, II, Open Court, 1916.

43 (X1.4) Bounded Functions

A. General Remarks

A complex-valued function defined on a subset E of the complex z -plane is called a **bounded function** defined on E if its range $f(E)$ is bounded, that is, if there exists a positive constant M such that $|f(z)| \leq M$ on E . However, when studying the theory of bounded functions, we usually restrict ourselves to the consideration of \dagger analytic or \dagger harmonic functions. On the other hand, the classes of functions $f(z)$ satisfying conditions such as $\operatorname{Re} f(z) > 0$ or $\alpha < \arg f(z) < \beta$ rather than $|f(z)| \leq M$ are studied by a method similar to that applied to the study of bounded functions.

\dagger Schwarz's lemma, \dagger Liouville's theorem, and \dagger Riemann's theorem on the removability of singularities (which will be explained later) are among the classical theorems in the theory of bounded functions.

B. Maximum Principle

When a function $f(z)$ is holomorphic and not constant in a domain D of the complex plane, $|f(z)|$ never attains its maximum in the interior of D . In particular, when $f(z)$ is continuous on the bounded closed domain $\bar{D} = D \cup \partial D$, the maximum of $|f(z)|$ on D is taken on its boundary ∂D . This fact is called the **maximum (modulus) principle**.

As a direct application of the maximum principle, we can deduce **Schwarz's lemma**: If a holomorphic function $f(z)$ in $|z| < R$ satisfies $|f(z)| \leq M$ and $f(0) = 0$, we have $|f(z)| \leq M \cdot |z|/R$ ($|z| < R$). The equality at z_0 , $0 < |z_0| < R$, occurs only for the functions $f(z) = e^{i\lambda} Mz/R$ (where A is a real constant).

C. Lindelöf's Principle

E. Lindelöf extended Schwarz's lemma and obtained various extensions of the maximum principle, from which he, together with E. Phragmén, deduced several useful theorems on the behavior of a function that is single-valued and holomorphic in a neighborhood of the boundary. We mention some representative theorems:

Let $z = \varphi(\zeta)$ and $w = \psi(\zeta)$ both be \dagger meromorphic and \dagger univalent functions in $|\zeta| < 1$ that map $|\zeta| < 1$ onto D_z and D_w , respectively. Set $\varphi(0) = z_0$ and $\psi(0) = w_0$. Let $D_z(\rho)$ and $D_w(\rho)$ denote the images of $|\zeta| \leq \rho$ ($0 < \rho < 1$) under the mappings φ and ψ , respectively. Under these circumstances, if a function $f(z)$ that is holomorphic in D_z satisfies $f(D_z) \subset D_w$ and $f(z_0) = w_0$, then $f(D_z(\rho)) \subset D_w(\rho)$. Furthermore, unless $f(z)$ maps D_z onto D_w univalently, $f(D_z(\rho))$ is contained in the interior of $D_w(\rho)$ (an extension of Schwarz's lemma).

Let $f(z)$ be analytic in a bounded domain D but not necessarily single-valued. Suppose that $|f(z)|$ is single-valued. Suppose, furthermore, that there is a positive constant M such that, for each boundary point ζ of D , except for a finite number of boundary points and for each $\varepsilon > 0$, the inequality $|f(z)| < M + \varepsilon$ holds on the intersection of D with a suitable neighborhood of ζ , and suppose also that each of the exceptional points has a neighborhood such that $f(z)$ is bounded on the intersection of D with this neighborhood. Under these assumptions we have $|f(z)| \leq M$. Moreover, if $|f(z_0)| = M$ at a point z_0 of D , then $f(z)$ is a constant (an extension of the maximum principle).

Let $f(z)$ be holomorphic in an angular domain W : $\arg z| < \alpha\pi/2$. Suppose that there is a constant M such that, for each $\varepsilon > 0$, each finite boundary point has a neighborhood such that $|f(z)| < M + \varepsilon$ on the intersection of D with this neighborhood, and that for some positive number $\beta > \alpha$ and for sufficiently large $|z|$ the inequality $|f(z)| < \exp|z|^{1/\beta}$ holds. Under these assumptions we have $|f(z)| \leq M$ in D (**Phragmén-Lindelöf theorem**).

Let $f(z)$ be a function that is holomorphic and bounded in a closed angular domain W : $\alpha \leq \arg z \leq \beta$ except for the point at infinity. Suppose that $f(z) \rightarrow a$ as $z \rightarrow \infty$ along a side of the angle and that $f(z) \rightarrow b$ as $z \rightarrow \infty$ along the other side of it. Then we have $a = b$ and $f(z) \rightarrow a$ uniformly as $z \rightarrow \infty$ in W (**Lindelöf's asymptotic value theorem**).

D. Bounded Functions in a Disk

If $f(z)$ is a bounded holomorphic function in the unit disk $|z| < 1$, it has a limit at every

point z_0 on the circle $C: |z|=1$, except for a set of 1-dimensional measure zero, as z tends to z_0 from within an angle with vertex at z_0 and contained in $|z|<1$ (or along a *Stolz's path at z_0) (**Fatou's theorem**). Under the same assumption, if the boundary value function $f(e^{i\theta}) = \lim_{r \rightarrow 1-0} f(re^{i\theta})$, which exists by Fatou's theorem, is equal to a constant a for a set of positive measure on the circle C , then $f(z) \equiv a$ in $|z|<1$ (**F. and M. Riesz theorem**). These theorems are valid for some kinds of meromorphic functions in $|z|<1$ (- 272 Meromorphic Functions D).

E. Three-Circle Theorem and Related Theorems

Let $f(z)$ be a function that is single-valued, holomorphic, and not identically zero in an annulus $\rho < |z| < R$. Set $M(r) \equiv \max_{|z|=r} |f(z)|$ ($\rho < r < R$) for $f(z)$. Then $\log M(r)$ is a \dagger convex function of $\log r$ in $\log \rho < \log r < \log R$ (**Hadamard's three-circle theorem**). The same assertion holds for a function $f(z)$ that is not necessarily single-valued, as long as $|f(z)|$ is single-valued. When $f(z)$ is single-valued, a stronger assertion can be obtained (O. Teichmüller). The following theorems are regarded as the analog of Hadamard's three-circle theorem for the respective basic regions:

Set $L(\sigma) \equiv \sup_{-\infty < t < \infty} |f(\sigma + it)|$ ($\alpha < \sigma < \beta$) for a function that is bounded and regular in a strip $\alpha < \operatorname{Re} z < \beta$. Then $\log L(\sigma)$ is a convex function of σ in $\alpha < \sigma < \beta$ (**Doetsch's three-line theorem**).

Set $I(\sigma) = \limsup_{t \rightarrow \infty} |f(\sigma + it)|$ ($\alpha < \sigma < \beta$) for a function that is holomorphic and bounded in a half-strip $\alpha < \operatorname{Re} z < \beta$, $\operatorname{Im} z > 0$. Then $\log I(\sigma)$ is a convex function of σ in $\alpha < \sigma < \beta$ (**Hardy-Littlewood theorem**).

Set

$$I_p(r) = \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^p d\theta$$

for a holomorphic function in a disk $|z| < R$. Then for every $p > 0$, $\log I_p(r)$ is an increasing convex function of $\log r$ for $-\infty < \log r < \log R$ (**Hardy's theorem**).

F. Hardy Class

Hardy's theorem motivates us to introduce a class of functions. An analytic function f in the unit disk is said to belong to the **Hardy class** H^p ($0 < p < \infty$) if $I_p(r, f) = (2\pi)^{-1} \int_0^{2\pi} |f(re^{i\theta})|^p d\theta$ remains bounded as $r \rightarrow 1$. For the case $p = \infty$, $f \in H^\infty$ if $I_\infty(r, f) = \max_{|z|=r} |f(z)|$ is bounded.

An analytic function f is said to be of the class N if $\int_0^{2\pi} \log^+ |f(re^{i\theta})| d\theta$ is bounded for $r < 1$. $f \in N$ if and only if $f = \varphi/\psi$, where $\varphi, \psi \in H^\infty$. $f \in H^p$ if and only if $|f|^p$ has a \dagger harmonic majorant. If $1 \leq p \leq \infty$, H^p is a \dagger Banach space with the norm $\|f\|_p = (\sup_{r < 1} I_p(r, f))^{1/p}$ for $p < \infty$ and $\|f\|_\infty = \sup_{r < 1} I_\infty(r, f)$ for $p = \infty$, and if $0 < p < 1$, H^p is a complete metric space with the metric $\sup_{r < 1} I_p(r, f - g)$. For $0 < p < q < \infty$ the inclusion relation $N \supset H^p \supset H^q \supset H^\infty$ holds, so that Fatou's theorem and that of F. and M. Riesz are valid for functions of H^p ($p > 0$).

For a nonnegative integer p and a sequence of nonzero complex numbers a_n , $|\alpha_n| < 1$, the infinite product $B(z) = z^p \prod_{n=1}^\infty (\bar{\alpha}_n(\alpha_n - z)/|\alpha_n| (1 - \bar{\alpha}_n z))$ converges locally uniformly in the unit disk if and only if $\sum_{n=1}^\infty (1 - |\alpha_n|) < \infty$. Then $B(z)$ is called a **Blaschke product** and $\{a_n\}$ a **Blaschke sequence**. If the sequence $\{a_n\} \cup \{0\}$ is the set of zeros of an analytic function f (an m -tuple zero appears m times in $\{\alpha_n\}$, and $z = 0$ is a p -tuple zero), $B(z)$ is said to be generated by the zeros of f . If $f \in N$, then $\sum_{n=1}^\infty (1 - |\alpha_n|) < \infty$ for the set $\{a_n\}$ of zeros off: An analytic function f in the unit disk is called an **inner function** if its modulus is less than 1 and its nontangential limit on the unit circle is of modulus 1 almost everywhere. A Blaschke product is an inner function. An inner function f without zeros is said to be singular if $f(0) > 0$. A **singular inner function** f can be represented by $f(z) = \exp \left\{ - \int_0^{2\pi} (e^{i\theta} + z)/(e^{i\theta} - z) d\mu(\theta) \right\}$ with a positive measure μ \dagger singular with respect to Lebesgue measure. An **outer function** $F(z)$ for H^p ($p > 0$) is an analytic function that can be represented by

$$F(z) = e^{ir} \exp \left\{ \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log \psi(\theta) d\theta \right\},$$

where r is a real number and $\psi(\theta)$ is a nonnegative function such that $\psi(\theta) \in L^p$ and $\log \psi(\theta) \in L^1$. An H^p -function f ($p > 0$) is factorized uniquely as $f(z) = B(z)S(z)F(z)$, where $B(z)$ is a Blaschke product (generated by the zeros off), $S(z)$ a singular inner function, and $F(z)$ is an outer function for H^p (with $\psi(\theta) = |f(e^{i\theta})|$) (**factorization theorem**). Conversely, such a product belongs to H^p . The **interpolation problem** raised by R. C. Buck asks whether or not there exists a sequence of distinct points $\{z_n \mid |z_n| < 1\}$ such that, given any bounded sequence of complex numbers $\{w_n\}$, there is a function $f \in H^\infty$ for which $f(z_n) = w_n$. The sequence $\{z_n\}$ in the problem is called an **interpolating sequence**. L. Carleson proved that a necessary and sufficient condition for $\{z_n\}$ to be an interpolating sequence is that there be a $\delta > 0$ such that $\prod_{n \neq k} |(z_n - z_k)/(1 - \bar{z}_k z_n)| \geq \delta$ ($k = 1, 2, \dots$). An interpolating sequence is a Blaschke sequence, and any sequence $\{z_n \mid |z_n| \rightarrow 1\}$ contains an interpolating subsequence

tending to the unit circle. An analogous theorem is obtained for a function of H^p ($p \geq 1$).

G. Corona Problem

The Hardy space H^∞ on the unit disk D is furthermore a commutative † Banach algebra with identity. Let M denote the † maximal ideal space of H^∞ . It is a compact Hausdorff space with respect to the † Gel'fand topology. To each point $z \in D$, there correspond a homomorphism $\varphi_z: \varphi_z(f) = f(z)$ for $f \in H^\infty$, and a maximal ideal $M_z = \{f \in H^\infty | f(z) = 0\}$ as its kernel. This correspondence gives a continuous injection of D into the maximal ideal space M of H^∞ . By identifying z and M_z , D is regarded as an open subset of M , and the **corona problem** then asks: Are there points of M outside the closure of D ? The negative answer was given by L. Carleson, that is, D is dense in A_4 (**corona theorem**). The corona theorem is obtained as a direct consequence of the following theorem and is equivalent to it. For functions $f_i \in H^\infty$, $i = 1, 2, \dots, n$, satisfying $\sum |f_i(z)| > \delta$ in D with some $\delta > 0$, there exist $g_i \in H^\infty$, $i = 1, \dots, n$, such that $\sum f_i(z)g_i(z) = 1$ in D . As a trivial consequence of the corona theorem, we have the following **cluster value theorem**: Let $C(f, \zeta)$ denote the † cluster set of $f \in H^\infty$ at ζ with $|\zeta| = 1$; then $C(f, \zeta) = \hat{f}(M_\zeta)$, where \hat{f} is the Gel'fand transform off and M_ζ is the † fiber of M over ζ . M is decomposed into pairwise disjoint † Gleason parts. A Gleason part is a connected open set or a singleton according as each point of it is captured in the closure of an interpolating sequence or not. Each point of the † Shilov boundary Γ of M forms by itself a trivial Gleason part, and Γ is † adherent to a sequence S in D with the property that the set of its nontangential limit points covers the unit circle (**Brown-Shield-Zeller theorem**). Let D be, in turn, a general bounded domain. The corona theorem is proved for some classes of such domains. A Shilov boundary Γ lies only over the set of points on ∂D which are not † removable for bounded analytic functions. It is adherent to a sequence S in D on which $\sup_S |f(z)| = \|f\|_\infty$. For any $f \in H^\infty$ and $\zeta \in \partial D$, the boundary of $\hat{f}(M_\zeta)$ is contained in $\hat{f}(\Gamma_\zeta)$ if $\Gamma_\zeta = \Gamma \cap M_\zeta$ is not empty. This, together with the cluster value theorem, has an important implication in the theory of cluster sets. The extremal problem of maximizing $f'(z_0)$, $z_0 \in D$, among all $f \in H^\infty$ satisfying $\|f\|_\infty \leq 1$ has the unique normalized solution G such that $G'(z_0) = \max |f'(z_0)|$. $G(z)$ is called the **Ahlfors function**. The Ahlfors function has unit modulus on the Shilov boundary (**Fisher's theorem**). The theory of H^∞ is also applied to the problem of

† rational approximation of bounded analytic functions.

H. Applications of the Maximum Principle

Theorems of the following type are useful for some problems of holomorphic functions:

Let D be a domain. Suppose that there exists an arc of angular measure α that is on a circle of radius R centered at a point z_0 of D and not contained in D . Let C denote the intersection of the boundary of D with the disk $|z - z_0| < R$. If $f(z)$ is a single-valued holomorphic function that satisfies $|f(z)| \leq M$, and if $\limsup_{z \rightarrow \zeta} |f(z)| \leq M$ for every $\zeta \in C$, then the inequality $|f(z_0)| \leq M^{1-1/n} m^{1/n}$ holds for every positive integer n satisfying $2\pi/n \leq \alpha$ (**Lindelöf's theorem**).

Let D be a domain bounded by two segments OA , OB both starting from 0 and making an angle $\pi\alpha$, and a Jordan arc \widehat{AB} , and let R be the maximal distance between 0 and the points on \widehat{AB} . Suppose that $f(z)$ is holomorphic in D and that $\limsup |f(z)|$ as $z \rightarrow \zeta \in \partial D$ with $z \in D$ is not greater than M for $\zeta \in OA \cup OB$ and m for $\zeta \in \widehat{AB}$. Then we have $|f(z)| \leq M^{1-\lambda} m^\lambda$ (where $\lambda = (|z|/R)^{1/\alpha}$) at every point on the bisector of the angle $\angle AOB$ in D (**Carleman's theorem**).

I. Holomorphic Functions with Positive Real Parts

Holomorphic functions with positive real parts are intimately connected with bounded functions. Concerning these functions we have the following classical result, which is equivalent to Schwarz's lemma: If $f(z)$ is holomorphic in $|z| < R$, $\operatorname{Re} f(z) \geq 0$ in the same domain, and $f(0) = 1$, then $(R - |z|)/(R + |z|) \leq \operatorname{Re} f(z) \leq (R + |z|)/(R - |z|)$ ($|z| < R$). The right or left inequality becomes equality for some z_0 , $0 < |z_0| < R$, only if $f(z) = (Rz_0 \mp z_0z)/(R \pm z_0z)$, respectively.

In order to prove various results for the class of functions with positive real part, **Herglotz's integral representation**, which is based on Poisson's integral representation and unique to this class, can be used effectively. It is given by

$$f(z) = \int_0^{2\pi} \frac{e^{i\varphi} + z}{e^{i\varphi} - z} d\rho(\varphi), \quad |z| < R,$$

where $\rho(\varphi)$ is monotone increasing (real-valued) with total variation 1 and is determined uniquely up to an additive constant by $f(z)$. An analogous integral representation is introduced for a holomorphic function in an annulus.

Similar to the case of holomorphic functions is the notion of Hardy classes to harmonic

functions, that is, a harmonic function u in the unit disk is said to be of the class h^p ($0 < p \leq \infty$) if $I_p(r, u)$ is bounded for $r < 1$. $u \in h^p$ if and only if $u + iv \in H^p$, where v is conjugate to u . h^1 is equal to the vector space of functions which are representable as a difference of two non-negative harmonic functions. The Herglotz theorem stated above is then a corollary of the theorem of integral representation for h^1 -functions.

J. Coefficient Problems

There are many classical results for partial sums and coefficients of the Taylor expansion of bounded functions in a disk. Let $f(z) = \sum_{n=0}^{\infty} c_n z^n$ be the Taylor expansion of a bounded function in $|z| < 1$. Set its partial sum $s_n(z) = \sum_{v=0}^n c_v z^v$ ($n=0, 1, \dots$), and let $t_n(z) = (1/(n+1)) \sum_{v=0}^n s_v(z)$ ($n=0, 1, \dots$), which is the sequence of the arithmetic means of the partial sums (the Fejér sums). Then $|f(z)| \leq 1$ in $|z| < 1$ if and only if $|t_n(z)| \leq 1$ for $|z| = 1$ ($n=0, 1, \dots$) (L. Fejér). Thus the sequences $\{t_n(z)\}$ for bounded functions $f(z)$ are uniformly bounded, whereas the sequences $\{s_n(z)\}$ are not uniformly bounded. Indeed the maximum value of $|s_n(1)|$ over the set of function f satisfying $|f(z)| \leq 1$ for $|z| < 1$ is

$$G_n = 1 + \sum_{j=1}^n \binom{-1/2}{j}^2 = \sum_{j=0}^n \left(\frac{1 \cdot 3 \cdot \dots \cdot (2j-1)}{2 \cdot 4 \cdot \dots \cdot 2j} \right)^2$$

($G_n \sim \pi^{-1} \log n$ as $n \rightarrow \infty$).

The following result is decisive for coefficient problems: Set $h_{\mu\nu} = \sum_{j=0}^{\mu} \bar{c}_{\mu-j} c_{\nu-j}$ ($\mu \leq \nu$), $h_{\nu\mu} = h_{\mu\nu}$ for $f(z) = \sum_{n=0}^{\infty} c_n z^n$, let m_n^2 be the maximal eigenvalue (a nonnegative real number) of the Hermitian matrix $(-h_{\mu\nu})_{\mu, \nu=0}^n$, and let $m = \lim m_n (\geq 0)$.

Let H be a Hermitian form in infinitely many variables given by

$$H = m^2 \sum_{\nu=0}^{\infty} \bar{x}_{\nu} x_{\nu} - \sum_{\mu, \nu=0}^{\infty} h_{\mu\nu} \bar{x}_{\mu} x_{\nu} = m^2 \sum_{\nu=0}^{\infty} |x_{\nu}|^2 - \sum_{\mu=0}^{\infty} \left| \sum_{\nu=0}^{\infty} c_{\nu} x_{\mu+\nu} \right|^2.$$

Then a necessary and sufficient condition for $|f(z)| \leq 1$ when $|z| < 1$ is that H be positive semidefinite, i.e., the sequence of the principal minor determinants of H ,

$$\Delta \begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix}, \quad m=1, 2, \dots,$$

is all positive or positive for an initial finite number of them and zero for the remainder (I. Schur).

A corresponding result for functions with positive real part in a disk can be stated in a simpler form: A holomorphic function $f(z) = 1/2 + \sum_{n=1}^{\infty} c_n z^n$ in $|z| < 1$ satisfies $\operatorname{Re} f(z) \geq 0$ if and only if

$$\begin{vmatrix} 1 & c_1 & \dots & c_n \\ \bar{c}_1 & 1 & \dots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & \bar{c}_{n-1} & \dots & 1 \end{vmatrix} \geq 0, \quad n=1, 2, \dots$$

(C. Carathéodory). Furthermore, for $n = 1, 2, \dots$, when we regard (c_1, \dots, c_n) as a point of complex n -dimensional Euclidean space, we can determine the domain of existence of points satisfying this criterion by Carathéodory. This result is generalized for coefficients of the Laurent expansion of a function that is holomorphic and single-valued in an annulus.

Next, there are some results for a function that omits two values: If $f(z) = a + a_1 z + \dots$ ($a_1 \neq 0$) is holomorphic in $|z| < R$ and $f(z) \neq 0, 1$ in the same domain $|z| < R$, then there exists a constant $L(a_0, a_1)$, depending only on a , and a_1 , such that $R \leq L(a_0, a_1)$ (Laudau's theorem). Under these circumstances, $|f(z)| \leq S(a_0, \theta)$ in $|z| \leq \theta R$ for $0 < \theta < 1$, where $S(a_0, 0)$ is a constant depending only on a , and θ (Schottky's theorem). These theorems have applications in value distribution theory. On the other hand, coefficient problems have been investigated as extremal problems in H^p -spaces under general conditions.

K. Angular Derivative

Let $f(z)$ be holomorphic in $|z| < 1$. If $f(z) \rightarrow w_0$ uniformly as $z \rightarrow z_0$ along Stolz paths with end point at z_0 and if the limit $\lim_{z \rightarrow z_0} ((f(z) - w_0)/(z - z_0)) \equiv D$ exists, we call D the **angular derivative** of $f(z)$ at z_0 . In the case of the half-plane $\operatorname{Re} z > 0$, the angular derivative at the point z_0 on the imaginary axis is similarly defined. It should be noted that $f(z) \rightarrow w_0$ is replaced by $1/f(z)$ for $w_0 = \infty$ and $1/(z - z_0)$ by z for $z_0 = \infty$. In the latter case, a Stolz path is a path contained in an angular domain $|\arg z| \leq \alpha$ ($\alpha < \pi/2$) and tending to ∞ . The study of angular derivatives was initiated by G. Julia (1920) and J. Wolff (1926) and was further advanced by Carathéodory (1929) and E. Landau and G. Varilon (1929).

A fundamental theorem for angular derivatives can be stated as follows: If a holomorphic function $f(z)$ in $\operatorname{Re} z > 0$ satisfies $\operatorname{Re} f(z) \geq 0$, there exists a constant c ($0 \leq c < +\infty$) such that $f(z)/z \rightarrow c$ and $f'(z) \rightarrow c$ uniformly as $z \rightarrow \infty$ along every Stolz path. Moreover, the p th derivative of $f(z)$ for an arbitrary positive p , denoted by $D^p f(z)$, has the property that $z^{p-1} D^p f(z) \rightarrow c/\Gamma(2-p)$ uniformly. Further-

Bounded Functions

more, the inequality $\operatorname{Re} f(z) \geq c \operatorname{Re} z$ holds everywhere in $\operatorname{Re} z > 0$. An analogous theorem is valid for the unit disk.

An important problem in the theory of conformal mapping is to find some condition for a mapping $w = f(z)$ of the unit disk (or half-plane) G onto a simply connected domain B to have a nonzero and finite angular derivative at a boundary point z_0 , that is, the condition for conformality at the boundary point. Carathéodory showed that a sufficient condition is the existence of two circles that are mutually inscribed and circumscribed at the boundary point $w_0 = f(z_0) \neq \infty$ of B and that lie inside and outside of B , respectively. Ahlfors later established a necessary and sufficient condition for the existence of the angular derivative by making use of his †distortion theorem for a strip domain. The angular derivative was used by Wolff in his research on the iteration of conformal mappings.

References

- [1] L. Bieberbach, *Lehrbuch der Funktionentheorie II*, Teubner, 1931 (Johnson Reprint Co., 1969).
- [2] C. Carathéodory, *Funktionentheorie I, II*, Birkhäuser, Basel, 1950; English translation, *Theory of functions*, Chelsea, I, 1958; II, 1960.
- [3] E. G. H. Landau, *Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie*, Springer, 1929.
- [4] J. E. Littlewood, *Lectures on the theory of functions I*, Clarendon Press, 1944.
- [5] P. L. Duren, *Theory of H^p -spaces*, Academic Press, 1970.
- [6] R. Nevanlinna, *Eindeutige Analytische Funktionen*, Springer, second edition, 1953; English translation, *Analytic functions*, Springer, 1970.
- [7] L. Carleson, *Interpolations by bounded analytic functions and the corona problem*, *Ann. Math.*, (2) 76 (1962).
- [S] H. S. Shapiro and A. L. Shield, *On some interpolation problem for analytic functions*, *Amer. J. Math.*, 83 (1961).
- [9] K. Hoffman, *Banach spaces of analytic functions*, Prentice-Hall, 1962.
- [10] K. Hoffman, *Bounded analytic functions and Gleason parts*, *Ann. Math.*, (2) 86 (1967).
- [11] M. L. Weiss, *Cluster sets of bounded analytic functions from a Banach algebraic viewpoint*, *Ann. Acad. Sci. Fenn.*, 367 (1965).
- [12] T. W. Gamelin, *Lectures on $H^\infty(D)$* , Univ. Nacional de la Plata, 1972.
- [13] T. W. Gamelin, *The algebra of bounded analytic functions*, *Bull. Amer. Math. Soc.*, 79 (1973).

44 (XVII.1 0) Branching Processes

A. General Remarks

A **branching process** is a mathematical model for random motion of a family of particles each of which is in an isolated process of multiplication and death. Examples of such random motions are population growth, miosis of genes, growth of the numbers of neutrons in an atomic chain reaction, and cascade showers of cosmic rays. The simplest and most fundamental branching process is discussed in Section B. For a historical introduction to the study of branching processes \rightarrow Kendall [2] and Harris [3].

B. Galton-Watson Processes

Although there is a similar process with continuous parameter $t \in \mathbf{R}$, we consider here only the case of a discrete time parameter ($i = 0, 1, 2, \dots$). Suppose that we are given a family of particles of the same kind. Each member of the family splits into several particles according to a given probability law independently of the other members and its own past history. Let Z_n be the number of particles of the family at a moment (or generation) n ; then $\{Z_n\}$ gives rise to a †Markov chain. This is called the **Galton-Watson process**. A precise mathematical description of the process is as follows: A Galton-Watson process is a Markov chain $\{Z_n, n = 0, 1, \dots\}$ on the nonnegative integers with †transition probability defined in terms of a given †probability distribution $\{p_k, k = 0, 1, \dots, p_k \geq 0, \sum_{k=0}^{\infty} p_k = 1$, by

$$p_{ij} = P[Z_{n+1} = j | Z_n = i] = \begin{cases} p_i^{*j} & \text{if } i \geq 1, j \geq 0, \\ \delta_{ij} & \text{if } i = 0, j \geq 0, \end{cases}$$

where the probability distribution $\{p_k^{*i}, k = 0, 1, \dots\}$ denotes the i -fold †convolution of the probability distribution $\{p_k, k = 0, 1, \dots\}$ and δ_{ij} is the †Kronecker delta. Then p_k is interpreted as the probability that an object existing in the n th generation has k children in the $(n+1)$ th generation. We assume that $Z_0 = 1$ a.s. for the rest of this section. An important tool in the analysis of the process is the †generating function $f(s) = \sum_{j=0}^{\infty} p_j s^j, |s| \leq 1$. Denote the generating function $\sum_{k=0}^{\infty} P(Z_n = k) s^k$ ($s \leq 1$) of Z_n by $f_n(s)$. Then we have $f_0(s) = s$, $f_1(s) = f(s)$, $f_{i+j}(s) = f_i(f_j(s))$ ($i, j = 0, 1, 2, \dots$), and the †expectation of Z_n is given by $E(Z_n) = m^n$, where $m = f'(1) = E(Z_1)$ is the expectation of Z_1 . A Galton-Watson process is said to be **subcritical**, **critical**, or **supercritical** if $m < 1$, $= 1$, or > 1 , respectively. If we have $Z_n =$

0 for some n , then $Z_{n+1} = Z_{n+2} = \dots = 0$. The probability $q = P(\lim_n Z_n = 0)$ is called the **extinction probability**. The case $f_1(s) \equiv s$ is excluded in the following. The extinction probability q of the process $\{Z_n\}$ is the smallest nonnegative solution of the equation $s = f(s)$. It is 1 if $m \leq 1$ and < 1 if $m > 1$. Moreover, $P(\lim_n Z_n = \infty) = 1 - q$ when $m > 1$, and hence the process $\{Z_n\}$ is \dagger transient.

If we put $W_n = Z_n/m^n$ when $m < \infty$, $\{W_n\}$ gives rise to a \dagger martingale, and the limit $W = \lim_{n \rightarrow \infty} W_n$ exists with probability 1. If $m > 1$, then the \dagger moment-generating function $\varphi(s) = E(\exp(-sW))$ of W satisfies the **Königs-Schröder equation** $\varphi(ms) = f(\varphi(s))$, $\text{Re } s \geq 0$. If $m \leq 1$, then $P(Z_n = 0 \text{ for some } n) = 1$. The following theorem on the conditional distribution of Z_n , given that $Z_n \neq 0$, was first proved by Yaglom [6] under moment restrictions. If $m < 1$, there exist $b_k = \lim_n P(Z_n = k | Z_n \neq 0)$, $k = 1, 2, \dots$. Then $\{b_k | k = 1, 2, \dots\}$ is a probability distribution and its generating function $g(s) = \sum_{k=1}^{\infty} b_k s^k$, $|s| \leq 1$, is the unique solution of the equation $g(f(s)) = mg(s) + 1 - m$, $|s| \leq 1$, among generating functions vanishing at 0. Furthermore, when $m < 1$, $\sum_{j=1}^{\infty} j b_j < \infty$ if and only if $\sum_{j=1}^{\infty} p_j(j \log j) < \infty$ (i.e., $E[Z_1 \log Z_1] < \infty$).

If $m = 1$ and $\sigma^2 = \text{variance of } Z_1 < \infty$, then $\lim_n P[Z_n/n > u | Z_n \neq 0] = \exp[-2u/\sigma^2]$, $u \geq 0$. If $\sigma^2 = \infty$, then this result is still correct with the interpretation that the limit in the left-hand side equals 1.

For further details regarding the theory of Galton-Watson processes \rightarrow [3, S].

C. Multitype Galton-Watson Processes

These are generalizations of the Galton-Watson process, involving k types of particles ($k \geq 2$), say, T_1, T_2, \dots, T_k . Let $p^i(r_1, r_2, \dots, r_k)$ be the probability that a particle of type T_i splits into r_m particles of type T_m ($m = 1, 2, \dots, k$), and set $f^i(s_1, s_2, \dots, s_k) = \sum_r p^i(r_1, r_2, \dots, r_k) s_1^{r_1} s_2^{r_2} \dots s_k^{r_k}$. The number of particles $Z_n = (Z_n^1, Z_n^2, \dots, Z_n^k)$ at a moment n gives rise to a \dagger Markov chain over Z_+^k , and its transition probability is determined by $p^i(r_1, r_2, \dots, r_k)$ as for Galton-Watson processes. Here Z_+^k denotes the space of all k -dimensional \dagger lattice points whose components are nonnegative integers. We now define a mathematical model of the process mentioned above. A Markov chain $\{Z_n | n = 0, 1, \dots\}$ on Z_+^k is called a **multi (k)-type Galton-Watson process** if its transition probability is given by

$$P(i, j) = P[Z_{n+1} = j | Z_n = i] = \text{coefficient of} \\ \prod_{i=1}^k s_i^{j_i} \text{ in } \prod_{l=1}^k (f^l(s_1, s_2, \dots, s_k))^{i_l}$$

for $i = (i_1, i_2, \dots, i_k), j = (j_1, j_2, \dots, j_k) \in Z_+^k, n = 0, 1, \dots$ and $s = (s_1, s_2, \dots, s_k) \in D_k$, where $D_k = \{s = (s_1, s_2, \dots, s_k) \in \mathbf{R}^k | |s_l| \leq 1, l = 1, 2, \dots, k\}$.

When $m_{ij} = \partial f^i(1, \dots, 1) / \partial s_j < \infty$ ($i, j = 1, 2, \dots, k$), the \dagger conditional expectation of Z_{n+m} , given Z_m , is given as $E[Z_{m+n} | Z_m] = Z_m M^n$, where M is the matrix (m_{ij}) . From now on we assume that $m_{ij} < \infty$ ($i, j = 1, 2, \dots, k$). A multi-type Galton-Watson process $\{Z_n | n = 0, 1, \dots\}$ is said to be singular if the generating functions $f^l(s_1, s_2, \dots, s_k)$, $l = 1, 2, \dots, k$, are all linear in s_1, s_2, \dots, s_k with no constant terms. We assume nonsingularity throughout. If there is a positive integer N such that every component of M^N is positive, the process Z_n is said to be **positively regular**. In this case, M has a positive \dagger eigenvalue λ that is simple, and $\lambda > |\mu|$ for all other eigenvalues μ . The eigenvalue λ plays the role of m in the case of the Galton-Watson process. For every i , $1 \leq i \leq k$, set $q^i = P[Z_n = 0 \text{ for some } n | Z_0 = e_i]$, where $e_i = (e_i^1, e_i^2, \dots, e_i^k)$, $e_i^j = \delta_{ij}$. Then q^i is called the extinction probability of the process $\{Z_n | n = 0, 1, \dots\}$ starting with a single particle of type T_i ; let $\mathbf{q} = (q^1, q^2, \dots, q^k)$. Suppose that the process $\{Z_n | n = 0, 1, \dots\}$ is positively regular and not singular. Then $\mathbf{q} = (1, 1, \dots, 1)$ if $\lambda \leq 1$, and $0 \leq q^i < 1$ ($i = 1, 2, \dots, k$) if $\lambda > 1$, and \mathbf{q} satisfies the equation $\mathbf{q} = (f^1(\mathbf{q}), f^2(\mathbf{q}), \dots, f^k(\mathbf{q}))$. Furthermore, the process $\{Z_n | n = 0, 1, \dots\}$ is \dagger transient, i.e., $P[Z_n = \mathbf{j} \text{ infinitely often}] = 0$ for any $\mathbf{j} \neq \mathbf{0}, \mathbf{j} \in Z_+^k$. Moreover, if $\lambda > 1$, then $W_n = Z_n/\lambda^n$ converges with probability 1, and

$$\lim_{n \rightarrow \infty} Z_n/\lambda^n = v W \quad \text{a.s.,}$$

where v denotes the positive left \dagger eigenvector $v = (v_1, \dots, v_k)$ for λ of the matrix M , and W is a nonnegative \dagger random variable. $P[W > 0] > 0$ holds if and only if

$$E[Z_{ij}^{(i)} \log Z_{ij}^{(i)}] < \infty \text{ for all } 1 \leq i, j \leq k. \quad (1)$$

Here $Z_{ij}^{(i)}$ denotes the number of type T_j particles in the first generation for a process with $Z_0 = e_i$, $1 \leq i \leq k$. If condition (1) holds, the \dagger moment-generating functions of W , $\varphi_i(\alpha) = E[\exp[-\alpha W] | Z_0 = e_i]$, $i = 1, 2, \dots, k$, $\alpha > 0$, satisfy $\varphi_i(\alpha) = f^i(\varphi_1(\alpha/\lambda), \varphi_2(\alpha/\lambda), \dots, \varphi_k(\alpha/\lambda))$, $1 \leq i \leq k$.

When $\lambda \leq 1$, a theorem analogous to Yaglom's holds. For further information \rightarrow Joffe and Spitzer (*J. Math. Anal. Appl.* 19 (1967)) and Athreya and Ney [S].

D. Markov Branching Processes

The branching processes dealt with in Sections B and C are limited in the sense that generation times are fixed. We now formulate a continuous-time version of branching processes. The treatment in this section is limited

to Markov processes that are extensions to continuous time of the Galton-Watson process. A Markov process $\{X_t, P_t\}$ on \mathbf{Z}_+^1 is called a **Markov branching process** if the transition probability $P(t, i, j) = P_t[X_t = j] (i, j \in \mathbf{Z}_+^1, t \geq 0)$ satisfies

$$\sum_{j=0}^{\infty} P(t, i, j) s^j = \left(\sum_{j=0}^{\infty} P(t, 1, j) s^j \right)^i$$

for all $i \in \mathbf{Z}_+^1$ and $|s| \leq 1$. Then there exist a positive constant a and a probability distribution $\{p_k; k = 0, 2, 3, \dots\}$ such that $\lim_{t \downarrow 0} (1 - P(t, i, i))/t = ia$ if $i \in \mathbf{Z}_+^1$ and $\lim_{t \downarrow 0} P(t, i, j)/t = iap_{j-i+1}$ if $j \geq i-1, j \neq i$. The +Kolmogorov equations are

$$\frac{d}{dt} P(t, i, j) = -jaP(t, i, j) + a \sum_{l=1, l \neq j}^{j+1} lp_{j-l+1} P(t, i, l) \quad (\text{forward equation})$$

and

$$\begin{aligned} \frac{d}{dt} P(t, i, j) &= -iaP(t, i, j) \\ &+ ia \sum_{l=i-1, l \neq i}^{\infty} p_{l-i+1} P(t, l, j) \end{aligned} \quad (\text{backward equation}).$$

Then a and $\{p_k; k = 0, 2, 3, \dots\}$ are interpreted as follows. An object existing at t has a probability adt of dying in the time interval $[t, t+dt]$ of length dt . If it dies at any time t , the probabilities are p_0, p_2, p_3, \dots that it is replaced by $0, 2, 3, \dots$ objects. When $\sum_{k=0}^{\infty} P(t, i, k) < 1$, the number of particles attains $+\infty$ in a finite time interval with positive probability. In order that $\sum_{k=0}^{\infty} P(t, i, k) = 1$ for all $i \in \mathbf{Z}_+^1$, it is necessary and sufficient that for each $1 > \varepsilon > 0, \int_{1-\varepsilon}^1 (f(u) - u)^{-1} du = \infty$, where $f(s) = \sum_{j=0, j \neq 1}^{\infty} p_j s^j$ for $|s| \leq 1$.

Most of the theory of the Galton-Watson process carries over to the continuous case. In particular, concerning the limit distribution of X_t , more precise results have been obtained. (For detailed discussion of Markov branching processes $\rightarrow [3, 8]$.)

Some work has also been done on the asymptotic behavior for temporally inhomogeneous cases of multitype Markov branching processes.

E. Branching Markov Processes

We now give a systematic treatment of general branching processes in which an object is characterized by a parameter x in a compact Hausdorff space S with a countable open base. The results are formulated in terms of strong Markov processes. Set $S^0 = \{\partial\}$, where

∂ is an extra point. For every positive integer n , let $S^n = \underbrace{S \times S \times \dots \times S}_n / \sim$, where \sim is the

*equivalence relation given by the *permutations of coordinates. S^n is compact with respect to the *quotient topology, and its *topological direct sum $S = \sum_{n=0}^{\infty} S^n$ is *locally compact. Let $\hat{S} = S \cup \{A\}$ be the one-point *compactification of S . A point $x \in \hat{S}$ is denoted by $\mathbf{x} = [x_1, x_2, \dots, x_n]$ if \mathbf{x} is the *equivalence class containing $(x_1, x_2, \dots, x_n) \in S^n$. $x = [x]$ is denoted simply by x . Set $\mathbf{B}(S) = \{f | f \text{ a bounded Borel measurable function on } S\}$, $\mathbf{B}^*(S) = \{f \in \mathbf{B}(S) | \|f\| < 1\}$, $\mathbf{C}(S) = \{f | f \text{ a bounded continuous function on } S\}$, and $\mathbf{C}^*(S) = \mathbf{C}(S) \cap \mathbf{B}^*(S)$. $\mathbf{B}(\hat{S}), \mathbf{C}(\hat{S}), \dots$ are defined similarly. For $f \in \mathbf{B}^*(S)$, define $\hat{f} \in \mathbf{B}(\hat{S})$ by $\hat{f}(\mathbf{x}) = 1$ if $\mathbf{x} = \partial, = \prod_{j=1}^n f(x_j)$ if $\mathbf{x} = [x_1, x_2, \dots, x_n] \in S^n$, and $= 0$ if $\mathbf{x} = A$. A *strong Markov process $X = \{X_t, \mathbf{P}_t\}$ on \hat{S} is called a **branching Markov process** if its semigroup $\{T_t\}$ (- 261 Markov Processes) satisfies

$$T_t \hat{f}(\mathbf{x}) = (\widehat{T_t f})_{\mathbf{S}}(\mathbf{x}) \quad \text{for every } \mathbf{x} \in \hat{S} \text{ and } f \in \mathbf{B}^*(S),$$

where for $g \in \mathbf{B}(\hat{S}), g|_S$ is the restriction of g on S [10-12]. We set (i) $Z_t = n$ if $X_t \in S^n, n = 0, 1, \dots, \infty$, where $S^\infty = \{A\}$, and (ii) $\tau = \inf\{t | Z_t \neq Z_0\}$. Z_t and τ are called the number of particles and the **first splitting time**, respectively. Let $X = \{x_t, P_x\}$ be a strong Markov process on S . We assume that the semigroup $\{H_t\}$ of X is a strongly continuous semigroup on $\mathbf{C}(S)$. Let φ_t be a nonnegative continuous *additive functional of X . Let $p_n(x), n = 0, 2, 3, \dots$ be a sequence of nonnegative functions in $\mathbf{B}(S)$ such that $\sum_{n=0, n \neq 1}^{\infty} p_n(x) = 1$. Consider a sequence of stochastic kernels $\pi_n(x, dy), n = 0, 2, 3, \dots$ on $S \times S^n$, i.e., for fixed $A \in \mathfrak{B}(S^n), \pi_n(x, A)$ is a measurable function of x , and for fixed x it is a probability on $(S^n, \mathfrak{B}(S^n))$, where $\mathfrak{B}(S^n)$ is the *topological σ -field on S^n . Set $\pi(x, D) = \sum_{n=0, n \neq 1}^{\infty} p_n(x) \pi_n(x, D \cap S^n)$ for $D \in \mathfrak{B}(\hat{S}), x \in S$, where $S(S)$ is the topological σ -field on \hat{S} . Then there exists a unique branching Markov process $X = \{X_t, \mathbf{P}_t\}$ on \hat{S} satisfying the following: (i) The Markov process $\{X_t, t < \tau, \mathbf{P}_x\}$ on S obtained from X by shortening its *lifetime is equivalent to the Markov process obtained by *killing X at a rate $d\varphi_t$ (- 261 Markov Processes). (ii) For any DE $S(S)$ and $\lambda > 0, E_x[\exp[-\lambda\tau]; X_t \in D | X_{t-}] = E_x[\exp[-\lambda\tau] | X_{t-}] \pi(X_{t-}, D)$ a.s. on $\{\tau < \infty\}$ for every $x \in S$. (iii) With probability one, $X_t = A$ for all $t \geq \lim_{n \uparrow \infty} \tau_n$, where τ_n denotes the n th splitting time [10]. This process $X = \{X_t, \mathbf{P}_t\}$ is referred to as the $\{x_t, \varphi_t, \pi\}$ -branching Markov process.

Example 1. Let $S = \{a_1, a_2, \dots, a_k\}$; then S can

be identified with \mathbf{Z}_+^k and $\hat{\mathbf{S}}$ with $\widehat{\mathbf{Z}}_+^k = \mathbf{Z}_+^k \cup \{\infty\}$. Therefore a branching Markov process X on $\hat{\mathbf{S}}$ is a Markov process on $\widehat{\mathbf{Z}}_+^k$ such that its transition probability $P(t, \mathbf{i}, \mathbf{j}), \mathbf{i}, \mathbf{j} \in \mathbf{Z}_+^k$ satisfies

$$\sum_{j_1, j_2, \dots, j_k \in \mathbf{Z}_+^k} P(t, \mathbf{i}, (j_1, j_2, \dots, j_k)) s_1^{i_1} s_2^{i_2} \dots s_k^{i_k} \\ = \prod_{l=1}^k \sum_{(j_1, j_2, \dots, j_k) \in \mathbf{Z}_+^k} P(t, e_l, (j_1, j_2, \dots, j_k)) \\ \times s_1^{j_1} s_2^{j_2} \dots s_k^{j_k} \Big)^{i_l},$$

$$P(t, \infty, \infty) = 1 \quad \text{for } \mathbf{i} = (i_1, i_2, \dots, i_k) \in \mathbf{Z}_+^k,$$

$$s = (s_1, s_2, \dots, s_k) \in D_k.$$

Then the process X is called a **multi (Q-type) Markov branching process**.

Example 2. Let $S = [0, \infty]$, $k(x)$ be a non-negative locally integrable function on $[0, \infty)$, and $p_n(x)$, $n = 0, 1, \dots$, be a sequence of non-negative measurable functions on $[0, \infty)$ such that $\sum_{n=0}^{\infty} p_n(x) = 1$ and $p_1(x) = 0$. We extend $k(x)$ and $p_n(x)$ ($n \neq 2$) as functions on $[0, \infty]$ by setting them to 0 at ∞ , and set $p_2(\infty) = 1$. Define a stochastic kernel $\pi(x, dy)$ on $S \times \hat{\mathbf{S}}$ by

$$\pi(x, D) = \begin{cases} \sum_{n=0}^{\infty} p_n(x) \delta_{\{[0, 0, \dots, 0]\}}^{(n)} (D \cap S^n) & \text{if } x \in [0, \infty), \\ \delta_{\{[\infty, \infty]\}} (D \cap S^2) & \text{if } x = \infty. \end{cases}$$

Let $X = \{x_t, P_x\}$ be the uniform motion on S with the semigroup $\{H_t\}$ such that for $f \in \mathbf{C}(S)$, $H_t f(x) = f(x+t)$ if $x \in [0, \infty)$ and $= f(\infty)$ if $x = \infty$. Set $\varphi_t = \int_0^t k(x_s) ds$. Then we have a $\{x_t, \varphi_t, n\}$ -branching Markov process, and we call it an **age-dependent branching process**.

We now discuss some fundamental properties of $\{x_t, \varphi_t, n\}$ -branching Markov processes. For the sake of simplicity we assume the following: (i) $\varphi_t = \int_0^t k(x_s) ds, k \in \mathbf{C}^+(S) = \{f \in \mathbf{C}(S) | f \geq 0\}$, (ii) $p_n(x) \in \mathbf{C}^+(S)$ and

$$\pi_n(x, D) = \delta_{\{[x, x, \dots, x]\}}(D)$$

for $D \in \mathcal{B}(S^n)$. Then $\{T_t\}$ is a strongly continuous semigroup on $\mathbf{C}(S) = \{f \in \mathbf{C}(\hat{\mathbf{S}}) | \lim_{x \rightarrow \infty} f(x) = 0\}$. Let A be the infinitesimal generator of $\{H_t\}$ (-261 Markov Processes) and $\mathcal{D}(A)$ be the domain of A . If $f \in \mathcal{D}(A) \cap \mathbf{C}^*(S)$, then $u(t, x) = (T_t \hat{f})_S(x) \in \mathcal{D}(A)$, and it satisfies $\partial u / \partial t = Au + k(F(\cdot; u) - u)$, $u(0+, x) = f(x)$, where $F(x; g) = \int_S g(y) \pi(x, dy)$ [10].

Let \mathcal{M} be the space of all nonnegative \dagger Radon measures on S and \mathcal{M}_0 be the subspace of \mathcal{M} consisting of all probability Radon measures on S endowed with the topology of weak convergence. Set $\mathcal{M} = \mathcal{M}_0 \cup \{\delta\}$, where δ is an extra point and $\tilde{\mathcal{M}} = [0, \infty] \times \mathcal{M}_0$. Define

a mapping $\rho: \tilde{\mathcal{M}} \ni (\bar{\lambda}, \lambda_0) \mapsto \rho(\bar{\lambda}, \lambda_0) \in \tilde{\mathcal{M}}$ by $\rho(\bar{\lambda}, \lambda_0) = \bar{\lambda} \lambda_0$ if $\lambda < \infty$ and $= \delta$ if $\lambda = \infty$, and define the topology of \mathcal{M} as the strongest of all the topologies rendering ρ continuous. Set $\mathbf{C}^{++}(S) = \{f \in \mathbf{C}(S) | f > 0\}$, and for $f \in \mathbf{C}^{++}(S)$, define a function $\Phi_f(\lambda)$ on $\tilde{\mathcal{M}}$ by $\Phi_f(\lambda) = \exp[-(\lambda, f)]$ if $\lambda \in \mathcal{M}$ and $= 0$ if $\lambda = \delta$, where $(\lambda, f) = \int_S f(x) \lambda(dx)$. Consider a mapping $\psi: \hat{\mathbf{S}} \rightarrow \tilde{\mathcal{M}}$ defined by $\psi(x) = 0$ if $x = \infty$, $= \sum_{k=1}^n \delta_{\{x_k\}}$ if $x = [x_1, x_2, \dots, x_n] \in S^n$, and $= \delta$ if $x = A$. Let \mathcal{M}_p be the subspace of $\tilde{\mathcal{M}}$ consisting of all non-negative integer-valued Radon measures on S , i.e., $\mathcal{M}_p = \{\psi(x) | x \in \hat{\mathbf{S}}\}$. We denote by \mathcal{D} the space of all right-continuous mappings $v: [0, \infty) \rightarrow \mathcal{M}_p$ whose discontinuities are at most of the first kind and such that $v_t = \delta$ for $t \geq \varsigma$ if $v_s = \delta$. Consider a branching Markov process $\{X_t, P_x\}$, and let P_{μ_x} be the probability law on \mathcal{D} of the stochastic process $\{v_t = \psi(X_t); 0 \leq t < \infty\}$ with $X_t = x$, where $\mu_x = \psi(x)$. Then the branching property can be rewritten as follows: For every $t \geq 0$, $f \in \mathbf{C}^{++}(S)$ and $\mu_1, \mu_2 \in \mathcal{M}_p$, $E_{\mu_1 + \mu_2}[\exp[-(v_t, f)]] = E_{\mu_1}[\exp[-(v_t, f)]] E_{\mu_2}[\exp[-(v_t, f)]]$ (\rightarrow Silverstein, Z. *Wahrscheinlichkeitstheorie und Verw. Gebiete*, 9 (1968); for a treatment of branching processes in the framework of \dagger martingale problems \rightarrow Holley and Stroock, *Publ. Inst. Math. Sci.*, 14 (1978)).

There are analogous processes which preserve the basic features of the branching property. A Markov process $\{\mu_t, P_\mu\}$ on $\tilde{\mathcal{M}}$ with δ as a \dagger trap is called a **continuous-state branching process** if it satisfies the property that for every $t \geq 0$, $f \in \mathbf{C}^{++}(S)$, and $\mu_1, \mu_2 \in \mathcal{M}$, $E_{\mu_1 + \mu_2}[\exp[-(\mu_t, f)]] = E_{\mu_1}[\exp[-(\mu_t, f)]] E_{\mu_2}[\exp[-(\mu_t, f)]]$. This concept was introduced by Jiiina [15] for some special cases. For further information \rightarrow Jifina (*Third Prague Conf.*, 1964), Lamperti (*Bull. Amer. Math. Soc.*, 73 (1967)), Watanabe (*J. Math. Kyoto Univ.*, 8 (1968)), and Fujimagari and Motoo (*Kôdai Math. Sem. Rep.*, 23 (1971)).

References

- [1] F. Galton and H. W. Watson, On the probability of the extinction of families, *J. Roy. Anthropol. Inst.*, 4 (1874), 138-144.
- [2] D. Kendall, Branching processes since 1873, *J. London Math. Soc.*, 41 (1966), 385-406.
- [3] T. E. Harris, The theory of branching processes, Springer, 1963.
- [4] M. S. Bartlett, An introduction to stochastic processes, Cambridge Univ. Press, 1966.
- [5] A. N. Kolmogorov and N. A. Dmitriev, Branching random processes (in Russian), *Dokl. Akad. Nauk SSSR*, (N.S.) 56 (1947), 5-8.

- [6] A. M. Yaglom, Some limiting theorems of branching random processes (in Russian), Dokl. Akad. Nauk SSSR, (N.S.) 56 (1947), 795–798.
- [7] H. Kesten and B. P. Stigum, A limit theorem for multidimensional Galton-Watson processes, Ann. Math. Statist., 37 (1966), 1211–1223.
- [8] K. B. Athreya and P. E. Ney, Branching processes, Springer, 1972.
- [9] B. A. Sevast'yanov, Branching processes (in Russian), Nauka, 1971.
- [10] N. Ikeda, M. Nagasawa, and S. Watanabe, Branching Markov processes I, II, III, J. Math. Kyoto Univ., 8 (1968), 233–278, 365–410; 9 (1969), 95–160; 11 (1971), 195–196.
- [11] J. E. Moyal, The general theory of stochastic population processes, Acta Math., 108 (1962), 1–31.
- [12] A. V. Skorokhod, Branching diffusion processes, Theory of Prob. Appl., 9 (1964), 445–449. (Original in Russian, 1964.)
- [13] T. W. Mullikin, Limiting distributions for critical multitype branching processes with discrete time, Trans. Amer. Math. Soc., 106 (1963), 469–494.
- [14] S. Watanabe, Limit theorem for a class of branching processes, Markov Processes and Potential Theory, J. Choiver (ed.), Wiley, 1967, 205–232.
- [15] M. Jiiina, Stochastic branching processes with continuous state space, Czech. Math. J., 8 (1958), 292–313.

45 (XVII.7) Brownian Motion

A. General Remarks

R. Brown, an English botanist, observed in 1827 that the minute particles comprising the pollen of plants, when suspended in water, exhibit peculiarly erratic movements [1]. The physical explanation of this phenomenon is that haphazard impulses are given to the suspended particles by collisions with molecules of the fluid. Let $X(t)$ be the x -coordinate of a particle at time t . Then $X(t)$ is treated as a random variable, and the distribution of $X(t) - X(s)$ is a normal distribution $N(0, D|t-s|)$, with mean 0 and variance $D|t-s|$, where D is a positive constant. To be more exact, such a family of random variables $\{X(t)\}$ is now considered as the family of random variables determining a stochastic process. Various aspects of the theory were analyzed by A. Einstein [2], L. Bachelier, N. Wiener, P. Lévy, and others.

B. Wiener Processes

Let T be the real line \mathbf{R}^1 or a subinterval. A stochastic process $\{X(t)\}_{t \in T}$, defined on a probability space $(\Omega, \mathfrak{B}, P)$ is called a **Wiener process** on \mathbf{R}^d if it satisfies the following three conditions: (1) $X(t, \omega) \in \mathbf{R}^d$ ($t \in T, \omega \in \Omega$). (2) The \mathbf{R}^d -valued random variables $X(t_1), X(t_j) - X(t_{j-1}), j = 2, 3, \dots, n$, are independent for any $t_1 < t_2 < \dots < t_n, t_j \in T$ ($j = 1, 2, \dots, n$), where n is an arbitrary positive integer. (3) If $X_i(t)$ is the i th component of the vector $X(t)$, then the $\{X_i(t)\}_{t \in T}$ ($1 \leq i \leq d$) are independent as stochastic processes and every increment $X_i(t) - X_i(s)$ is normally distributed with mean 0 and variance $|t-s|$. A Wiener process is also called a **Brownian motion**. A Wiener process is a temporary homogeneous additive process. A separable Wiener process has continuous paths with probability 1. Conversely, if $\{X(t)\}_{t \in T}$ is a temporary homogeneous additive process on \mathbf{R}^1 whose sample function is continuous with probability 1 and the increment $X(t) - X(s)$ has mean 0 and variance $|t-s|$, then it is a Wiener process (-407 Stochastic Processes, 406 Stochastic Differential Equations). Let $\{\mathcal{F}_t\}_{t \geq 0}$ be an increasing family of σ -subfields of \mathfrak{B} . We assume that $\{\mathcal{F}_t\}$ is right continuous. A d -dimensional continuous process $X = (X(t))_{t \geq 0}$, is called a d -dimensional $\{\mathcal{F}_t\}$ -**Brownian motion** if it is $\{\mathcal{F}_t\}$ -adapted and satisfies

$$E[\exp\{i\langle \xi, X(t) - X(s) \rangle\} | \mathcal{F}_s] \\ = \exp[-(t-s)|\xi|^2/2]$$

a.s. for every $\xi \in \mathbf{R}^d, 0 \leq s < t$,

where $|\xi|$ denotes the norm of $\xi \in \mathbf{R}^d$ [9]. Then X satisfies conditions (1)–(3) mentioned above, and hence it is a Brownian motion on \mathbf{R}^d (-406 Stochastic Differential Equations).

Let $\{X_k(\omega)\}$ ($k = 0, 1, \dots$) be a sequence of independent random variables defined on a probability space $(\Omega, \mathfrak{B}, P)$ such that each X_k has the normal distribution $N(0, 1)$ with mean 0 and variance 1. Then the series

$$\frac{t}{\sqrt{\pi}} X_0(\omega) + \sqrt{\frac{2}{\pi}} \sum_{n=1}^{\infty} X_n(\omega) \frac{\sin nt}{n}, \quad t \in [0, \pi] \equiv T,$$

converges uniformly in $t \in T$ with probability 1, and its limit, denoted by $X(t, \omega)$, is a Wiener process [3, 10].

Let $W^d = C([0, \infty) \rightarrow \mathbf{R}^d)$ be the space of all continuous functions $w: [0, \infty) \rightarrow \mathbf{R}^d$ endowed with the topology of uniform convergence on finite intervals and $\mathfrak{B}(W^d)$ be the topological σ -field. Let $X = (X(t))_{t \geq 0}$ be a Brownian motion on \mathbf{R}^d and μ be the probability law of $X(0)$. The probability law P^X of the Brownian motion X on $(W^d, \mathfrak{B}(W^d))$ is called the d -

dimensional Wiener measure with the initial distribution μ [9, 11, 12]. For a probability μ on $(\mathbf{R}^d, \mathfrak{B}(\mathbf{R}^d))$ the d -dimensional Wiener measure P_μ with the initial distribution μ is a probability on $(W^d, \mathfrak{B}(W^d))$ characterized by the property that for every $0 = t_0 < t_1 < \dots < t_n$ and $A_j \in \mathfrak{B}(\mathbf{R}^d)$, $j = 1, 2, \dots, n$,

$$P_\mu[w(t_1) \in A_1, w(t_2) \in A_2, \dots, w(t_n) \in A_n] \\ = \int_{\mathbf{R}^d} \int_{A_1} \dots \int_{A_n} \prod_{j=1}^n \left(\frac{1}{\sqrt{2\pi(t_j - t_{j-1})}} \right)^d \\ \times \exp \left[-\frac{|x_j - x_{j-1}|^2}{2(t_j - t_{j-1})} \right] \mu(dx_0) dx_1 \dots dx_n,$$

where $\mathfrak{B}(\mathbf{R}^d)$ denotes the topological σ -field on \mathbf{R}^d .

C. Brownian Motion as a Diffusion Process

In terms of the general framework of \dagger Markov processes, Brownian motion is a typical example of a diffusion process (\rightarrow 115 Diffusion Processes). Let $X = \{X_t(w), \mathbf{R}^d, P_x\}$ be a continuous Markov process on \mathbf{R}^d (\rightarrow 261 Markov Processes) with the \dagger transition probability

$$P(t, x, B) = \int_B (2\pi t)^{-d/2} e^{-\frac{1}{2t}|x-y|^2} dy, \\ t > 0, \quad x \in \mathbf{R}^d, \quad B \in \mathfrak{B}(\mathbf{R}^d).$$

For each $x \in \mathbf{R}^d$, the process $\{X_t(w), P_x\}$ is a Wiener process in the sense mentioned above. The Markov process X , which is a collection of Wiener processes $\{X_t, P_x\}$ starting at x , is said to be a **d-dimensional Brownian motion**. A d -dimensional Brownian motion possesses the \dagger strong Markov property. Let \mathfrak{G} be the \dagger generator of the \dagger semigroup T_t corresponding to X . A bounded uniformly continuous function f defined on \mathbf{R}^d belongs to the domain of \mathfrak{G} if its partial derivatives $\partial f / \partial x_i$ and $\partial^2 f / \partial x_i \partial x_j$, $i, j = 1, 2, \dots, d$, are bounded and uniformly continuous. For such a function f we have $\mathfrak{G}f(x) = (1/2)\Delta f(x)$, where Δ is the \dagger Laplacian [8, 12].

D. Brownian Motions and Potentials

For $\alpha > 0$, the function $G_\alpha(x)$ defined by

$$G_\alpha(x) = \frac{1}{2} \int_0^\infty e^{-\alpha t} (2\pi t)^{-d/2} \exp \left(-\frac{1}{2t}|x|^2 \right) dt$$

is said to be the **cc-order Green's function**.

Since Brownian motion is \dagger nonrecurrent for $d \geq 3$, the limit $G_{0+}(x) = \lim_{\alpha \downarrow 0} G_\alpha(x)$, $x \in \mathbf{R}^d$, exists for $d \geq 3$, and $G_{0+}(x)$ is equal to $K_\alpha(x) = (\Gamma(d/2 - 1)/4\pi^{d/2})|x|^{-d+2}$, which is the kernel for the \dagger Newtonian potentials. Brownian motion is \dagger recurrent when $d \leq 2$ and $G_{0+}(x) = +\infty$, $x \in \mathbf{R}^d$. In this case, $K_\alpha(x)$ is defined by

$K_0(x) = \lim_{\alpha \downarrow 0} (G_\alpha(x) - G_\alpha(x_0))$, and $K_0(x) = (1/2\pi) \log(1/|x|)$, when $d = 2$ and $|x_0| = 1$.

This is the kernel for the \dagger logarithmic potentials. When $d = 1$ and $x_0 = 0$, $K_\alpha(x) = -|x|/2$. Using this relationship, we can express many concepts of classical potential theory in an elegant form in probability language. Let $X = \{X(t), P_x\}$ be the d -dimensional Brownian motion. Given a set A , set $\sigma_A = \inf\{t \geq 0, X(t) \in A\}$, where the infimum over the empty set is understood to be $+\infty$. Then σ_A is called the \dagger hitting time of X for the set A (\rightarrow 261 Markov Processes). For a Green domain (i.e., a domain which is a \dagger Green space) D in \mathbf{R}^d ($d \geq 2$), set $2g^D(t, x, y)dy = P_x(X(t) \in dy, \sigma_{\partial D} > t)$, $x, y \in D$. Then the right-hand side of this equation is the \dagger transition probability of the Brownian motion on D with the \dagger absorbing barrier ∂D . Then $G^D(x, y) = \int_0^\infty g^D(t, x, y)dt$ is \dagger Green's function of D . If B is a compact subset of a Green domain D or an open subset with compact closure $\bar{B} \subset D$, then the hitting probability $p(x) = P_x(\sigma_B < \sigma_{\partial D})$ is the equilibrium potential of B relative to D .

Suppose A is an \dagger analytic subset of \mathbf{R}^d . \dagger Blumenthal's O-I law implies that $P_x(\sigma_A = 0) = 1$ or 0 . The point x is said to be **regular** for A if this probability is 1 and **irregular** for A otherwise. Let B be a compact subset in \mathbf{R}^d ($d \geq 2$). Then \dagger Wiener's test (\rightarrow 120 Dirichlet Problem) states that x is regular or irregular for B according as the following series diverge or converge:

$$\sum_{k=1}^\infty 2^{k(d-2)} C(B_k), \quad d \geq 3, \\ \sum_{k=1}^\infty k C(B_k), \quad d = 2,$$

where $C(B_k)$ is the \dagger Newtonian capacity (the logarithmic capacity relative to a bounded domain when $d = 2$) of the set $B_k = \{y \mid 2^{-(k+1)} \leq |y-x| < 2^{-k}\} \cap B$ [12]. Suppose D is a bounded domain in \mathbf{R}^d ($d \geq 2$). A point $x \in \partial D$ is regular or irregular for $\mathbf{R}^d - D$ according as x is regular or irregular in the sense of the \dagger Dirichlet problem for D . Given a continuous boundary function f on ∂D , $u(x) = E_x(f(X(\sigma_{\partial D}))$ is the solution of the generalized Dirichlet problem for D . Given $x \in D$, the distribution $h(x, B) = P_x(X(\sigma_{\partial D}) \in B)$ ($B \in \mathfrak{B}(\partial D)$) used in the solution $u(x) = \int_{\partial D} f(y)h(x, dy)$ of the generalized Dirichlet problem is the \dagger harmonic measure of ∂D as viewed from x .

E. 1-Dimensional Brownian Motion

In his monograph [6], P. Lévy gave a profound description of the fine structure of the individual 1-dimensional Brownian path

$\{X(t, \omega)\}_{t \geq 0}$. Let us set

σ_a = the hitting time to the point a in \mathbf{R}' ,

$$m(t, \omega) = \min_{\sigma_0 \leq s \leq t} X(s, \omega),$$

$$M(t, \omega) = \max_{\sigma_0 \leq s \leq t} X(s, \omega),$$

$$Y_0(t, \omega) = |X(t, \omega)|,$$

$$Y_1(t, \omega) = \begin{cases} X(t, \omega), & t < \sigma_0(\omega), \\ M(t, \omega) - X(t, \omega), & t \geq \sigma_0(\omega), \end{cases}$$

and

$$Y_2(t, \omega) = \begin{cases} X(t, \omega), & t < \sigma_0(\omega), \\ X(t, \omega) - m(t, \omega), & t \geq \sigma_0(\omega). \end{cases}$$

Then we have (1) $P_0(M(t) > a) = 2P_0(M(t) > a, X(t) > a) = 2P_0(M(t) > a, X(t) < a) = P_0(|X(t)| > a)$ (reflection principle of D.André).

(2) The stochastic process $\{\sigma_a, 0 \leq a < \infty, P_0\}$ is a †one-sided stable process with exponent $1/2$, that is, it is additive and homogeneous with the law

$$\begin{aligned} P_0(\sigma_b - \sigma_a \leq t) &= P_0(\sigma_{b-a} \leq t) \\ &= \int_0^t \frac{b-a}{\sqrt{2\pi s^3}} e^{-(b-a)^2/2s} ds, \\ &0 \leq a < b, \quad t \geq 0. \end{aligned}$$

(3) Let $\varphi^{-1}(t, \omega)$ be the right continuous inverse function of

$$\varphi(t, \omega) = \int_0^t \chi_{[0, \infty)}(X(s, \omega)) ds,$$

where $\chi_{[0, \infty)}(\cdot)$ is the †indicator function of the interval $[0, \infty)$. Set $Y_3(t, \omega) = X(\varphi^{-1}(t, \omega), \omega)$. Then the four processes $\{Y_i(t, \omega), 0 \leq t < \infty, P_x\}$ ($x \in [0, \infty), 0 \leq i \leq 3$) on $[0, \infty)$ have the same probability law. Each of them is a diffusion process with transition probability

$$\begin{aligned} P(t, x, B) &= \int_B \frac{1}{\sqrt{2\pi t}} (e^{-|x-y|^2/2t} + e^{-|x+y|^2/2t}) dy, \\ B &\in \mathfrak{B}([0, \infty)), \quad t > 0, \quad x \in [0, \infty), \end{aligned}$$

and is said to be a Brownian motion on $[0, \infty)$ with a †reflecting barrier at the origin. (4) As a consequence of (3), if $X(0, \omega) = 0$ a.s., then, for fixed t , $M(t)$, $-m(t)$, and $Y_i(t)$ ($0 \leq i \leq 3$) have a common distribution. For example,

$$\begin{aligned} P_0(M(t) \geq a) &= P_0(\sigma_a \leq t) = 2P_0(X(t) > a) \\ &= \sqrt{\frac{2}{\pi t}} \int_a^\infty e^{-x^2/2t} dx \quad a > 0, \\ P_0(X(t) \in da, M(t) \in db) &= \sqrt{\frac{2}{\pi t^3}} (2b-a) e^{-(2b-a)^2/2t} da db, \quad 0 \leq a < b. \end{aligned}$$

(5) The diffusion process $\{X(t, \omega), 0 \leq t < \sigma_0, P_x\}$ ($x \in (0, \infty)$) obtained from a 1-dimensional Brownian motion by shortening its †lifetime is called a Brownian motion on $(0, \infty)$ with an †absorbing barrier at the origin, and its transition probability is given by

$$\begin{aligned} P(t, x, B) &= \int_B \frac{1}{\sqrt{2\pi t}} (e^{-|x-y|^2/2t} - e^{-|x+y|^2/2t}) dy, \\ B &\in \mathfrak{B}([0, \infty)), \quad t > 0, \quad x \in (0, \infty). \end{aligned}$$

The **arcsin law** is valid for many functionals of 1-dimensional Brownian motion. For example,

$$P_0\left(\int_0^t \chi_{[0, \infty)}(X(s, \omega)) ds \leq \theta\right) = \frac{2}{\pi} \arcsin \sqrt{\frac{\theta}{t}}, \quad 0 < \theta \leq t,$$

$$P_0(\tau_t(\omega) \leq s) = (2/\pi) \arcsin \sqrt{s/t}, \quad 0 < s \leq t,$$

where $\tau_t(\omega) = \sup\{s | X(s, \omega) = 0, 0 \leq s \leq t\}$.

The visiting set $\mathcal{Z}(\omega) = \{t | X(t, \omega) = 0\}$ of a Brownian path is a †totally disconnected set. Its †Lebesgue measure is 0, and the †Hausdorff-Besikovich dimension number of $\mathcal{Z}(\omega)$ is $1/2$.

Consider a 1-dimensional Wiener process $\{X(t)\}_{0 \leq t < \infty}$ starting at the origin. Let us set

$$B_1(t) = \begin{cases} 0 & t = 0, \\ tX(1/t) & t > 0, \end{cases}$$

and

$$B_2(t) = cX(t/c^2), \quad t \geq 0, \quad (c > 0).$$

Then the stochastic processes $\{B_1(t)\}_{0 \leq t < \infty}$ and $\{B_2(t)\}_{0 \leq t < \infty}$ are likewise 1-dimensional Wiener processes starting at the origin. Hence the properties of the Wiener process starting at the origin in a neighborhood of $t = 0$ ($t = \infty$) can be obtained from those in a neighborhood of $t = \infty$ ($t = 0$).

F. d-Dimensional Brownian Motion

Almost all paths of d-dimensional Brownian motion are continuous but are not of †bounded variation on any finite interval. Accordingly, they cannot have lengths. A positive, continuous, increasing function φ defined on $[t_0, \infty)$ with $t_0 > 0$ is said to **belong to the upper (lower) class with respect to local continuity** if $P_0(\{(\inf\{t | |X(t, \omega)| > \sqrt{t} \varphi(1/t), t > 0\}) > 0\}) = 1$ (0). Kolmogorov's test states that φ belongs to the upper class or to the lower class with respect to local continuity according as

$$\int_{t_0}^{\infty} t^{-1} (\varphi(t))^d e^{-\varphi^2(t)/2} dt$$

converges or diverges. For example,

$$\varphi(t) = (2 \log_{(2)} t + (d+2) \log_{(3)} t + 2 \log_{(4)} t + \\ + 2 \log_{(n-1)} t + (2+\delta) \log_{(n)} t)^{1/2}$$

belongs to the upper or lower class with respect to local continuity according as $\delta > 0$ or $\delta \leq 0$, where $\log_{(n)} t = \log \log_{(n-1)} t$ and $\log_{(1)} t = \log t$. The law of the iterated logarithm for 1-dimensional Brownian motion,

$$P_0 \left(\limsup_{t \rightarrow s} \frac{|X(t, \omega) - X(s, \omega)|}{\sqrt{2|t-s| \log \log 1/|t-s|}} = 1 \right) = 1,$$

is a special case of this example. For the case $d = 1$, consider the **space-time Brownian motion** $\{(-t, X(t)), P_0\}$. Set $D_\varphi = \{(s, x) | -1/t_0 \leq s \leq 0, \sqrt{-s} \varphi(-1/s) \leq x < \infty\}$. Then $O = (O, O)$ is a regular point of D_φ or an irregular point of D_φ for the space-time Brownian motion according as φ belongs to the lower or upper class with respect to local continuity. (- 261 Markov Processes). Thus Kolmogorov's test is the Wiener test for space-time Brownian motion. Let ψ be a positive, continuous decreasing function defined on $[t_0, \infty)$ with $t_0 > 0$. Then

$$P_0 \left(\left(\inf \{t | |X(t, \omega)| < \sqrt{t} \psi(1/t), t > 0\} \right) > 0 \right)$$

= 0 or 1.

We have

$$\int_{t_0}^{\infty} \frac{1}{t} (\psi(t))^{d-2} dt = \infty \text{ or } < \infty \quad \text{when } d \geq 3,$$

and

$$\int_{t_0}^{\infty} \frac{dt}{t |\log \psi(t)|} = \infty \text{ or } < \infty \quad \text{when } d = 2.$$

To describe uniform continuity of a path on the interval $[0, 1]$, take a positive, continuous increasing function ψ defined on $[t_0, \infty)$ with $t_0 > 0$ and set $\varphi(t) = \sqrt{t} \psi(1/t)$. Then ψ is said to **belong to the upper class with respect to uniform continuity** if almost all paths $X(t, \omega)$ ($0 \leq t \leq 1$) satisfy the \dagger Lipschitz condition relative to φ , that is, for almost all ω there exists an $\varepsilon(\omega) > 0$ such that $0 < t-s < \varepsilon(\omega)$ implies $|X(t, \omega) - X(s, \omega)| < \varphi(|t-s|)$. And ψ is said to **belong to the lower class with respect to uniform continuity** if almost all paths $X(t, \omega)$ ($0 \leq t \leq 1$) do not satisfy the Lipschitz condition relative to φ . Then ψ belongs to the upper or lower class with respect to uniform continuity according as

$$\int_{t_0}^{\infty} \psi(t)^{d+2} e^{-\psi^2(t)/2} dt$$

converges or diverges [16]. For example,

$$\psi(t) = (2 \log t + (d+4) \log_{(2)} t + 2 \log_{(3)} t + \\ + 2 \log_{(n-1)} t + (2+\delta) \log_{(n)} t)^{1/2}$$

belongs to the upper or lower class with respect to uniform continuity according as $\delta > 0$ or $\delta \leq 0$. The following theorem on the uniform continuity of 1-dimensional Brownian motion is a special case of this criterion:

$$P_0 \left(\limsup_{\substack{t-s \rightarrow 0 \\ 0 \leq t, s \leq 1}} \frac{|X(t, \omega) - X(s, \omega)|}{\sqrt{2|t-s| \log 1/|t-s|}} = 1 \right) = 1.$$

Now we state some other properties of Brownian paths. Let A be a set of zero \dagger outer capacity in \mathbf{R}^d ($d \geq 2$). Then $P_x(X(t) \in A \text{ for some } t > 0) = 0$ for any $x \in \mathbf{R}^d$. Let A be a plane set with positive \dagger inner capacity. Then $P_x(X(t) \in A \text{ for infinitely many } t \text{ larger than any given } s > 0) = 1$ for any $x \in \mathbf{R}^2$. With probability 1, the Lebesgue measure of the set $\{X(t, \omega) | 0 \leq t < \infty\}$ in \mathbf{R}^d is zero for $d \geq 2$. With probability 1 this set is everywhere dense in \mathbf{R}^2 when $d = 2$ and nowhere dense in \mathbf{R}^d when $d \geq 3$ [13, 15, 17]. Almost all 2-dimensional Brownian paths have k -fold multiple points for any integer $k \geq 2$. In the 3-dimensional case, almost all paths have \dagger double points but cannot have any triple point. In the $d(\geq 4)$ -dimensional case, almost all paths have no double point [18-20].

G. Itô's Formula and Brownian Local Time

The following formula of Itô [21] is of fundamental importance in the theory of stochastic processes.

Itô's formula. Let $X = \{X(t) = (X_1(t), X_2(t), \dots, X_d(t)), P_x\}$ be a d -dimensional Brownian motion and $f(x) = f(x_1, x_2, \dots, x_d)$ be a C^2 -function defined on \mathbf{R}^d . Let us set $\langle p(t) = f(X(t)) - f(X(0))$. Then $\{\varphi(t)\}_{t \geq 0}$ is a continuous \dagger additive functional of the Brownian motion X (- 261 Markov Processes), and we have

$$f(X(t)) - f(X(0)) = \sum_{i=1}^d \int_0^t \left(\frac{\partial}{\partial x_i} f \right) (X(s)) dX_i(s) \\ + \frac{1}{2} \int_0^t (\Delta f)(X(s)) ds, \quad t \geq 0,$$

where the first term in the right-hand side is a \dagger stochastic integral with respect to the Brownian motion X [22] (- 406 Stochastic Differential Equations). In particular, if f is a C^3 -function defined on \mathbf{R}^d , Itô's formula can be rewritten as

$$f(X(t)) - f(X(0)) = \sum_{i=1}^d \int_0^t \left(\frac{\partial}{\partial x_i} f \right) (X(s)) \circ dX_i(s), \quad t \geq 0,$$

where the right-hand side is a stochastic integral of the \dagger Stratonovich type with respect to

the Brownian motion X [9] (- 406 Stochastic Differential Equations).

Consider a 1-dimensional Brownian motion $X = (X(t), P_x)$. By the **local time** or the **sojourn time density** of X we mean a family of non-negative random variables $\{\varphi(t, x)\}$ ($t \in [0, \infty)$, $x \in \mathbf{R}^1$) such that, with probability 1, the following holds:

- (a) the mapping $[0, \infty) \times \mathbf{R}^1 \ni (t, x) \mapsto \varphi(t, x) \in \mathbf{R}^1$ is continuous,
- (b) for every Borel subset A of \mathbf{R}^1 and $t \geq 0$,

$$\int_0^t \chi_A(X(s)) ds = 2 \int_A \varphi(t, x) dx.$$

The notion of the local time of Brownian motion was first introduced by Lévy [6]. The family of random variables $\{\varphi(t, x)\}$ ($t \geq 0$, $x \in \mathbf{R}^1$) defined by

$$\varphi(t, x) = (X(t) - x)^+ - (X(0) - x)^+$$

$$- \int_0^t \chi_{(x, \infty)}(X(s)) dX(s)$$

satisfies properties (a) and (b) mentioned above and hence it is a local time [9, 22]. Here a^+ is the bigger of a and 0. It is clear that the local time $\varphi(t, x)$ of X is given by

$$\varphi(t, x) = \lim_{\varepsilon \downarrow 0} \frac{1}{4\varepsilon} \int_0^t \chi_{(x-\varepsilon, x+\varepsilon)}(X(s)) ds$$

for every $x \in \mathbf{R}^1$, $t \geq 0$.

Furthermore, we have

$$P_0 \left(\limsup_{b-a=\delta \downarrow 0, a < b} \frac{|\varphi(t, b) - \varphi(t, a)|}{\sqrt{\delta \log 1/\delta}} \leq 2 \sqrt{\max_{a \in \mathbf{R}^1} \varphi(t, a)} \right) = 1,$$

and

$$P_0 \left(\limsup_{\delta \downarrow 0} \frac{|\varphi(t, \delta) - \varphi(t, 0)|}{\sqrt{\delta \log 1/\delta}} \leq 2 \sqrt{\varphi(t, 0)} \right) = 1.$$

Lévy [6] studied the fine structure of the local time. Define the visiting sets $\mathcal{Z}^+ = \{t \mid Y_0(t, \omega) = 0\}$ and $\mathcal{Z}^- = \{t \mid Y_1(t, \omega) = 0\}$, where $Y_0(t, \omega)$ and $Y_1(t, \omega)$ are the stochastic processes defined in Section E above. Then we have

$$P_0 \left(\lim_{\varepsilon \downarrow 0} \sqrt{\frac{\pi \varepsilon}{2}} \times [\text{the number of flat stretches of } M(s, \omega) (0 \leq s \leq t) \text{ of length } \geq \varepsilon] = M(t, \omega), t \geq 0 \right) = 1.$$

Since the two diffusion processes $X^+ = \{Y_0(t), 0 \leq t < \infty, P_x\}$ and $X^- = \{Y_1(t), 0 \leq t < \infty, P_x\}$ define the same probability law on $W([0, \infty))$, there exists a functional $\varphi^+(t, \omega)$ of X^+ corresponding to $M(t, \omega)$ of X^- . Here,

$W([0, \infty))$ denotes the space of all continuous functions $w: [0, \infty) \rightarrow [0, \infty)$. The flat stretches of the graph of $\varphi^+(t, \omega)$ are the open intervals \mathcal{Z}_n ($n \geq 1$) such that $\sum_{n \geq 1} \mathcal{Z}_n = [0, \infty) - \mathcal{Z}^+$. Then we have

$$P_0 \left(\lim_{\varepsilon \downarrow 0} \sqrt{\frac{\pi \varepsilon}{2}} \times [\text{the number of intervals } \mathcal{Z}_n \subset [0, t) \text{ of length } \geq \varepsilon] = \varphi^+(t, \omega), t \geq 0 \right) = 1.$$

Furthermore, we have

$$P_0 \left(\lim_{\varepsilon \downarrow 0} \sqrt{\frac{\pi}{2\varepsilon}} \times [\text{the total length of the intervals } \mathcal{Z}_n \subset [0, t) \text{ of length } < \varepsilon] = \varphi^+(t, \omega), t \geq 0 \right) = 1$$

and

$$P_0 \left(\lim_{\varepsilon \downarrow 0} (2\varepsilon)^{-1} \int_0^t \chi_{[0, \varepsilon)}(Y_0(s, \omega)) ds = \varphi^+(t, \omega), t \geq 0 \right) = 1.$$

Let $d_\varepsilon(t, \omega)$ be the number of times that the reflecting Brownian path $Y_0(s, \omega)$ crosses down from $\varepsilon > 0$ to 0 before time t . Then

$$P_0 \left(\lim_{\varepsilon \downarrow 0} \varepsilon d_\varepsilon(t, \omega) = \varphi^+(t, \omega), t \geq 0 \right) = 1.$$

H. Flows and Random Distributions

The flow derived from the 1-dimensional Wiener process $\{X_s\}_{-\infty < t < \infty}$ is \dagger Kolmogorov's flow. It has \dagger mixing properties of all orders and is \dagger ergodic (- 136 Ergodic Theory, 39.5 Stationary Processes). The \dagger stationary process with independent values at every point corresponding to the \dagger characteristic functional

$$\exp \left(-\frac{1}{2} \int_{-\infty}^{\infty} \varphi(t)^2 dt \right)$$

on the \dagger Schwartz space \mathcal{Y} defines the same probability law with the stationary process obtained by differentiation of the Wiener process in the \dagger distribution sense [23] (- 395 Stationary Processes, 407 Stochastic Processes).

I. Generalizations of Brownian Motion

In addition to the Brownian motion described above, there are several stochastic processes that are also called Brownian motion. A Gaussian system $\{X(a)\}_{a \in \mathbf{R}^n}$ defined on a probability space $(\Omega, \mathfrak{B}, P)$ is said to be a **Brownian motion** with an N -dimensional time parameter

if (i) $E(X(a))=0$, (ii) $E(X(a)X(b))=\frac{1}{2}(|a|+|b|-|a-b|)$, (iii) $P(X(0)=0)=1$. Let a^* be the spherical inversion of $a \in \mathbf{R}^N$ with respect to the unit sphere. Set $X^*(a)=|a|X(a^*)$ ($a \neq 0$) and $X^*(0)=0$. Then $\{X^*(a)\}_{a \in \mathbf{R}^N}$ defines the same probability law with a Brownian motion with an N -dimensional time parameter. Almost all paths of a Brownian motion with an N -dimensional time parameter are continuous. A positive, continuous, increasing function φ defined on $[t_0, \infty)$ with $t_0 > 0$ is said to belong to the upper (lower) class with respect to local continuity if the probability that the closure of the set $\{a \mid |X(a, \omega)| > \sqrt{|a|} \varphi(1/|a|), |a| > 0\}$ contains the origin 0 is equal to 0 (1). Then φ belongs to the upper or lower class with respect to local continuity according as the integral

$$\int_{t_0}^{\infty} \frac{1}{t} (\varphi(t))^{2N-1} e^{-\varphi^2(t)/2} dt$$

converges or diverges. For example,

$$\varphi(t) = (2 \log_{(2)} t + (2N-1) \log_{(3)} t + 2 \log_{(4)} t + 2 \log_{(n-1)} t + (2+\delta) \log_{(n)} t)^{1/2}$$

belongs to the upper or lower class with respect to local continuity according as $\delta > 0$ or $\delta \leq 0$. As a special case, we have

$$P\left(\limsup_{a \rightarrow 0} \frac{|X(a, \omega)|}{\sqrt{2|a| \log \log 1/|a|}} = 1\right) = 1.$$

Take a positive, continuous increasing function φ defined on $[t_0, \infty)$ with $t_0 > 0$, and set $\psi(t) = \sqrt{t} \varphi(1/t)$. If almost all paths $X(a, \omega)$ ($|a| \leq 1$) satisfy the Lipschitz condition relative to $\psi(t)$, then φ is said to belong to the upper class with respect to uniform continuity. It is said to belong to the lower class with respect to uniform continuity if, with probability 1, these paths do not satisfy the Lipschitz condition relative to $\psi(t)$. Then φ belongs to the upper or lower class with respect to uniform continuity according as

$$\int_{t_0}^{\infty} t^{N-1} (\varphi(t))^{4N-1} e^{-\varphi^2(t)/2} dt$$

converges or diverges (T. Sirao, *Nagoya Math. J.*, 17 (1960)). For example,

$$\varphi(t) = (2N \log t + (4N+1) \log_{(2)} t + 2 \log_{(3)} t + \dots + 2 \log_{(n-1)} t + (2+\delta) \log_{(n)} t)^{1/2}$$

belongs to the upper or lower class with respect to uniform continuity according as $\delta > 0$ or $\delta \leq 0$. As a special case,

$$P\left(\limsup_{\substack{|a-b| \rightarrow 0, |a|, |b| \leq 1}} \frac{|X(a, \omega) - X(b, \omega)|}{\sqrt{2N|a-b| \log 1/|a-b|}} = 1\right) = 1$$

(For general information about Brownian motion with a multidimensional time parameter \rightarrow P. Lévy [6] and H. P. McKean, *Theory of Prob. Appl.* 8 (1963).)

Let $X(t)$ be a Wiener process. L. S. Ornstein and G. E. Uhlenbeck based their investigation of the irregular movements of small particles immersed in a liquid on Langevin's equation

$$dU(t) = -\alpha U(t)dt + \beta dX(t),$$

where $U(t)$ is the velocity of a particle. The first term on the right-hand side of this †stochastic differential equation (- 406 Stochastic Differential Equations) is due to frictional resistance or its analog, which is thought to be proportional to the velocity. The second term represents random external force. The stationary solution of this equation is given by

$$U(t) = \int_{-\infty}^t \beta e^{-\alpha(t-u)} dX(u).$$

The stochastic process $\{U(t)\}_{-\infty < t < \infty}$ is a stationary †Gaussian Markov process with covariance function $\gamma(t) = (\beta^2/2\alpha)e^{-\alpha|t|}$. This process is called Ornstein-Uhlenbeck Brownian motion [24] (\rightarrow 176 Gaussian Processes).

Brownian motion has been defined on state spaces that are †Riemannian spaces or +Lie groups, and its properties are being investigated (\rightarrow 5 Additive Processes, 115 Diffusion Processes).

References

- [1] R. Brown, A brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; On the general existence of active molecules in organic and inorganic bodies, *Philos. Mag., Ann. Philos. New Series*, 4 (1828).
- [2] A. Einstein, Investigations on the theory of the Brownian movement, Methuen, 1926.
- [3] R. E. A. C. Paley and N. Wiener, Fourier transforms in the complex domain, Amer. Math. Soc. Colloq. Publ., 1934.
- [4] N. Wiener, Differential space, *J. Math. and Phys.*, 2 (1923), 131–174.
- [5] P. Lévy, Le mouvement brownien plan, *Amer. J. Math.*, 62 (1940), 487–550.
- [6] P. Lévy, Processus stochastiques et mouvement brownien, Gauthier-Villars, 1948.
- [7] J. L. Doob, Stochastic processes, Wiley, 1953.
- [8] K. Itô, Lectures on stochastic processes, Tata Inst, 1960.
- [9] N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, North-Holland and Kodansha, 1981.

- [10] K. Itô and M. Nisio, On the convergence of sums of independent Banach space-valued random variables, Osaka J. Math., 5 (1968), 35-48.
- [11] P. Billingsley, Convergence of probability measure, Wiley, 1963.
- [12] K. Itô and H. P. McKean, Jr., Diffusion processes and their sample paths, Springer, 1965.
- [13] J. L. Doob, Semi-martingales and subharmonic functions, Trans. Amer. Math. Soc., 77 (1954), 86-121.
- [14] A. Ya. Khinchin (Khintchin) Asymptotische Gesetze der Wahrscheinlichkeitsrechnung, Erg. Math., Springer, 1933 (Chelsea, 1948).
- [15] S. Kakutani, Two-dimensional Brownian motion and harmonic functions, Proc. Imp. Acad. Tokyo, 20 (1944), 706-714.
- [16] K. L. Chung, P. Erdos, and T. Sirao, On the Lipschitz's condition for Brownian motion, J. Math. Soc. Japan, 11 (1959), 263-274.
- [17] S. Kakutani, On Brownian motions in n -space, Proc. Imp. Acad. Tokyo, 20 (1944), 648-652.
- [18] A. Dvoretzky, P. Erdos, and S. Kakutani, Double points of Brownian motion in n -space, Acta Sci. Math. Szeged., 12 (1950), 75-81.
- [19] A. Dvoretzky, P. Erdos, and S. Kakutani, Multiple points of paths of Brownian motion in the plane, Bull. Res. Council Israel, 3 (1954), 364-371.
- [20] A. Dvoretzky, P. Erdos, S. Kakutani, and S. J. Taylor, Triple points of Brownian paths in 3-space, Proc. Cambridge Philos. Soc., 53 (1957), 856-862.
- [21] K. Itô, On a formula concerning stochastic differentials, Nagoya Math. J., 3 (1951), 55-65.
- [22] H. P. McKean, Stochastic integrals, Academic Press, 1972.
- [23] T. Hida, Brownian motion, Springer, 1980.
- [24] G. E. Uhlenbeck and L. S. Ornstein, On the theory of Brownian motion, Phys. Rev., 36 (1930), 823-841.

C

46 (X.32) Calculus of Variations

A. General Remarks

One of the first objects of differential calculus was to systematize the theory of the extrema of functions of a finite number of independent variables. In the calculus of variations we consider **functionals** (i.e., real-valued or complex-valued functions defined on a function space $\{u\}$ consisting of functions defined on a certain domain B), originally so named by J. Hadamard. Such a functional is denoted by $J[u]$ (or simply $J[u]$). A function u , which is considered an independent variable of the functional, is called an **argument function** (or **admissible function**).

Concrete examples of functionals are the length of a curve $y = f(x)$ and the area of a surface $z = z(x, y)$, which are expressed by

$$L[y] = \int_{x_0}^{x_1} \sqrt{1 + (y'(x))^2} dx$$

and

$$S[z] = \iint_B \sqrt{1 + z_x^2 + z_y^2} dx dy,$$

respectively. Furthermore, consider a curve $y = y(x)$ connecting two given points (x_0, y_0) and (x_1, y_1) with $y_1 > y_0$. The time in which a particle slides down without friction from (x_0, y_0) to (x_1, y_1) along this curve under constant gravity acting in the direction of the positive y -axis is expressed by the functional

$$J[y] = k \int_{x_0}^{x_1} \sqrt{(1 + y'^2)/(y - y_0)} dx, \quad k \text{ constant. (1)}$$

Let $F(\dots)$ be a known real-valued function that depends on a certain number of independent variables and on argument functions of these variables, together with derivatives of these functions up to a certain order. Then a typical problem in the calculus of variations is formulated as an extremal problem of a functional that is expressed by the integral with $F(\dots)$ as the integrand, for example,

$$J[u] = \int_{x_0}^{x_1} F(x, u(x), u'(x), \dots, u^{(m)}(x)) dx,$$

$$J[u, v] = \iint_{S \text{ or } B} F(x, y, u(x, y), u_x, u_y, \dots, v(x, y), v_x, v_y, \dots) dx dy.$$

For instance, a curve minimizing (1) is the **curve of steepest descent**.

In extremal problems, suitable boundary conditions can be assigned to argument func-

tions. On the other hand, there are so-called **conditional problems in the calculus of variations**. A typical example of this sort is the **isoperimetric problem**, i.e., the determination of the curve that bounds a domain with maximal area among all curves on a plane with given length. In general, an extremal problem of a functional under a subsidiary condition that the value of another given functional remain fixed is called a **generalized isoperimetric problem**. In addition to these, there are **Lagrange's problem**, in which a finiteness condition is imposed, and **Hilbert's problem**, in which a condition consisting of differential equations is imposed.

The birth of the calculus of variations was almost simultaneous with that of differential and integral calculus. Johann Bernoulli, Jakob Bernoulli, L. Euler, and others had dealt with several concrete problems of the calculus of variations when in 1760, J. L. Lagrange introduced a general method of dealing with variational problems connected with mechanics. Then an equation bearing the name of Euler or Lagrange was introduced.

B. Euler's Equation

As an example, consider the simplest variational problem

$$J[y] = \int_{x_0}^{x_1} F(x, y(x), y'(x)) dx = \min. \quad (2)$$

Let the boundary condition $y(x_0) = y_0, y(x_1) = y_1$ be assigned to the argument function $y(x)$. Consider a family of admissible functions $Y(x; \varepsilon) = y(x) + \varepsilon \eta(x)$, where $\eta(x)$ is any fixed function vanishing at both endpoints and ε is a parameter. If $y(x)$ gives the minimum of $J[y]$, then the function of $\varepsilon, J[Y]$, must attain a minimum for $\varepsilon = 0$. The condition $(\partial J[Y]/\partial \varepsilon)_{\varepsilon=0} = 0$ is written in the form

$$0 = \int_{x_0}^{x_1} (F_y \eta + F_{y'} \eta') dx \\ = \int_{x_0}^{x_1} \eta \left(F_y - \frac{d}{dx} F_{y'} \right) dx$$

by taking into account the boundary condition. By making use of the arbitrariness of $\eta(x)$, we conclude that

$$0 = F_y - \frac{d}{dx} F_{y'} = F_y - F_{y'x} - y' F_{y'y} - y'' F_{y'y'} \quad (3)$$

holds, in view of the following lemma: Let $\varphi(x)$ be continuous in $[x_0, x_1]$ and $\eta(x)$ be a function of class C^p that satisfies $\eta(x_0) = \eta(x_1) = \dots = \eta^{(q)}(x_0) = 0, \eta(x_1) = \eta'(x_1) = \dots = \eta^{(q)}(x_1) = 0$ ($0 \leq q \leq p$). If $\int_{x_0}^{x_1} \eta(x) \varphi(x) dx = 0$ holds for any such $\eta(x)$, then $\varphi(x) = 0$. (Here class C^p can be

replaced by class C^ω when q is supposed finite. If $\eta \in C^\omega$, then $q = \infty$ is admitted.) This is called the **fundamental lemma in the calculus of variations**. We call (3) the **Euler-Lagrange differential equation** (or **Euler's equation** for the extremal problem). Since this equation is of the second order, $y(x)$ can be determined by means of the boundary condition.

The quantity $[F]_y = F_y - \frac{d}{dx} F_{y'}$ contained in equation (3) is called the **variational derivative** of F with respect to y . Furthermore, $\delta y = \eta d\epsilon$ and $\delta J = (\partial J / \partial Y) \delta Y$, $d\epsilon$ are called the **first variations** of the argument function y and of the functional $J[y]$, respectively. If $\eta(x)$ is not subject to the condition that it must vanish at the endpoints, then the **first variation** of J becomes

$$\delta J = \int_{x_0}^{x_1} [F]_y \delta y dx + [F_{y'} \delta y]_{x_0}^{x_1}.$$

In comparison with an ordinary extremal problem $f(x_1, \dots, x_n) = \min$ in differential calculus, $[F]_y$ and δJ correspond to df , respectively. In general, a solution of the Euler-Lagrange differential equation $[F]_y = 0$ is called a **stationary function** for the variational problem, and its graph is called a **stationary curve**.

For a variational problem involving several argument functions, we have only to write the system of Euler-Lagrange differential equations corresponding to them. For a problem

$$J[y] = \int_{x_0}^{x_1} F(x, y, y', \dots, y^{(m)}) dx = \min$$

whose integrand involves derivatives of higher orders of an argument function, the Euler-Lagrange differential equation is

$$[F]_y = \sum_{\mu=0}^m (-1)^\mu \frac{d^\mu}{dx^\mu} F_{y^{(\mu)}} = 0.$$

For a problem involving a double integral

$$J[u] = \iint_{S_B} F(x, y, u, u_x, y_y) dx dy = \min,$$

the equation is

$$0 = [F]_u = F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y}.$$

For a generalized isoperimetric problem, for example $J[y] = \min$, $K[y] = c$, the Euler-Lagrange equation becomes

$$[F + \lambda G]_y = 0$$

with the so-called **Lagrange multiplier** λ , where F and G denote the integrands of J and K , respectively. Two integration constants contained in a general solution of this differential equation of the second order and an undeter-

mined constant λ can be determined, for instance, by the boundary condition $y(x_0) = y_0$, $y(x_1) = y_1$, and a subsidiary condition $K[y] = c$. As an example, for the classical proper isoperimetric problem $F = y$, $G = \sqrt{1 + y'^2}$, the equation is $1 - \lambda(y' / \sqrt{1 + y'^2})' = 0$, which after integration leads to $(x - \alpha)^2 + (y - \beta)^2 = \lambda^2$ (\rightarrow 228 Isoperimetric Problems).

Besides the case of fixed endpoints, there is a boundary condition, for instance, that an endpoint (x_1, y_1) of the argument function $y = y(x)$ must lie on a given curve $T(x, y) = 0$. For the case of such a movable endpoint, the extremal function is subject to the **condition of transversality**,

$$(F - y' F_{y'}) T_y - F_{y'} T_x = 0, \quad x = x_1.$$

C. Sufficient Conditions

A. M. Legendre introduced the notion of the second variation, corresponding to differential quotients of the second order in differential calculus, in order to discuss **sufficient conditions**. Concerning the simplest problem (2), the inequality $F_{y'y'}(x, y_0(x), y'_0(x)) \geq 0$ is necessary in order for $y_0(x)$ to give the minimum. Conversely, the inequality $F_{y'y'} > 0$ and Jacobi's condition (which is stated below) imply that $y = y_0(x)$ gives a **weak minimum**. Here "weak minimum" means the minimum when a family of admissible functions $\{y - y_0, < \epsilon, y' - y'_0 < \epsilon\}$ is considered a neighborhood of y_0 .

Jacobi's condition: Let u be a solution of a linear ordinary differential equation of the second order,

$$\frac{d}{dx} \left(F_{y'y'} \frac{du}{dx} \right) - \left(F_{yy} - \frac{d}{dx} F_{yy'} \right) u = 0,$$

$$u(x_0) = 0;$$

then the smallest zero of u that is greater than x_0 (i.e., the **conjugate point** of x_0) is greater than the right endpoint x_1 .

K. Weierstrass derived sufficient conditions for a strong minimum by extending the range of admissible functions to $\{y - y_0 < \epsilon\}$. Results that were obtained until about that time constitute the content of what is usually called the **classical theory of the calculus of variations**.

If for the variations problem (2) there exists a unique curve through every point in a domain on the xy -plane that belongs to a one-parameter family of stationary curves of the functional

$$J[y] = \int_{x_0}^{x_1} F(x, y, y') dx,$$

then the domain is called a **field** of stationary curves. Let the parameter value of the curve through a point (x, y) in such a family of

stationary curves $y = \varphi(x; a)$ be denoted by $\alpha = \alpha(x, y)$. The slope $p(x, y) = [\varphi'(x; \alpha)]_{x=\alpha(x, y)}$ is called the slope of the field at the point (x, y) or, by regarding x, y as variables, the **slope function** of the field. The value of a curvilinear integral

$$I_C = I[y] = \int_C (F(x, y, p) - (y' - p)F_y(x, y, p)) dx$$

is then determined and depends only on the two endpoints of the curve C . We call I_C **Hilbert's invariant integral**. In view of the property mentioned above, we can denote the value of the functional J for a function $y = y(x)$ representing a curve C by J_C . Then any admissible curve C that passes through a field embedding a stationary curve C_0 satisfies

$$0 \leq \Delta J = J_C - J_{C_0} = \int_C \mathcal{E}(x, y; p, y') dx.$$

Here

$$\begin{aligned} \mathcal{E}(x, y; p, y') \\ = F(x, y, y') - F(x, y, p) - (y' - p)F_y(x, y, p) \end{aligned}$$

is the \mathcal{E} -function introduced by Weierstrass. For C_0 to give the minimum of $J[y]$, it suffices that $\mathcal{E} \geq 0$ hold for every point (x, y) in the field and every value y' (\rightarrow 279 Morse Theory).

D. Optimal Control

Let a system of differential equations

$$\begin{aligned} dx_i/dt &= f_i(x_1, \dots, x_n; u_1, \dots, u_k), \\ (u_1, \dots, u_k) &\in \Omega; \quad x_i(t_0) = x_i, \quad i = 1, \dots, n, \end{aligned} \quad (4)$$

be given, where u_1, \dots, u_k are parameters. In general, a problem of **optimal control** is to determine $u_j = u_j(t)$ ($t_0 \leq t \leq t_1$) such that the value of a functional

$$J[u] = \int_{t_0}^{t_1} F(x_1, \dots, x_n; u_1, \dots, u_k) dt$$

assumes a minimum, where $x_i(t)$ are the solutions of (4) and are considered functions in u_1, \dots, u_k , and t . Such a problem is a kind of conditional variational problem. But since the existence region of u is restricted, certain conditions in the form of inequalities are imposed, and furthermore u is not necessarily continuous, and in many cases the problem cannot be treated within the classical theory of the calculus of variations (\rightarrow 86 Control Theory).

E. The Direct Method in the Calculus of Variations

In mathematical physics \dagger variational principles are derived from discussions of formal corre-

spondence between a functional $J[u]$ to be minimized and Euler's equation for $J[u]$. This is certainly one of the important methods in the calculus of variations, but it is also possible to investigate a stationary function u_0 on the basis of its stationary character and independently of Euler's equation. This is called the **direct method** in the calculus of variations. It plays an important role in the theoretical treatment of the existence and uniqueness of solutions, and it is also significant as a technique for approximate or numerical solutions. When a differential equation is given independently of the calculus of variations, it is possible to apply the direct method if a functional whose Euler's equation is the given differential equation can be constructed.

Let D be a bounded domain in m -dimensional space and $f \in L_2(D)$ be a real-valued function. Consider the variational problem of minimizing the functional

$$J[u] = \int_D |\text{grad } u|^2 dx - 2 \int_D f u dx.$$

Here we suppose the set of admissible functions, denoted by A_J , to be the Hilbert space obtained by completing the function space $C_0^\infty(D)$ with respect to the norm

$$N(u) = \left(\int_D |\text{grad } u|^2 dx + \int_D u^2 dx \right)^{1/2}$$

Utilizing F. Riesz's representation theorem in Hilbert spaces, it can be shown that there exists a minimum value l in A , which is uniquely realized by certain $u_0 \in \bar{A}_J$. Since the function u_0 belongs to A , it can be shown that the boundary condition

$$u|_{\partial D} = 0 \quad (5)$$

is satisfied in a generalized sense. Furthermore, in view of $J[u_0] \leq J[u_0 + \varphi]$ being valid for any $\varphi \in C_0^\infty(D)$, it can be verified that the equation

$$-\Delta u = f \quad (6)$$

is satisfied in D in the sense of differentiation of \dagger distributions. In other words, the stationary function u_0 is a solution in the wide sense (a tweak solution) of the classical boundary value problem for \dagger Poisson's equation formulated by (6) and (5). If a function space A , with $\bar{A}_J \supset A_J \supset C_0^\infty(D)$ is taken as the set of admissible functions, the value $l = J[u_0]$ becomes the greatest lower bound of J in A_J .

In this case, if $\{u_n\}_{n=1}^\infty$ is any **minimizing sequence** from A_J , that is, if $u_n \in A$, $n = 1, 2, \dots$; $J[u_n] \rightarrow l$ ($n \rightarrow \infty$), then it converges to u_0 in the sense that

$$N(u_n - u_0) \rightarrow 0 \quad (n \rightarrow \infty). \quad (7)$$

In other words, the solution in the wide sense

u_0 of the boundary value problem can be constructed as the limit of a minimizing sequence that consists of sufficiently smooth functions vanishing on the boundary. In the proof of the fact that a solution in the wide sense u_0 coincides with the classical solution of the boundary value problem under an assumption of suitable smoothness for f and ∂D , a standard argument has been established for proving the regularity of a solution in the wide sense.

The technique of obtaining the solution of the boundary value problem as the limit of a minimizing sequence was proposed by B. Riemann concerning the classical [†]Dirichlet problem and was completed by D. Hilbert. This pioneering work led to the recent treatment of boundary value problems by utilizing Hilbert spaces. For ⁺Self-adjoint boundary value problems, the method stated for the above example has been generalized almost directly to the cases of differential operators of higher order and with variable coefficients (-323 Partial Differential Equations of Elliptic Type). By making use of some auxiliary arguments, this technique can be applied extensively to the construction of several kinds of mapping functions in the theory of functions of a complex variable, to the solution of [†]integral equations of the second kind, and also in other fields [3, 4].

The eigenvalue problem, which is formulated by

$$Hu = \lambda u, \quad u \neq 0, \quad (8)$$

with a ⁺Self-adjoint operator H in a Hilbert space, can also be transformed into a variational problem for the [†]Rayleigh quotient

$$R[u] = (Hu, u) / \|u\|^2 \quad (9)$$

(→ 298 Numerical Computation of Eigenvalues).

F. Solution of Differential Equations by the Direct Method

In view of the convergence shown in (7), a minimizing sequence can be regarded as an approximating sequence for a solution of the boundary value problem or a stationary function u_0 . Let a function

$$u_n = u_n(x; c) = u_n(x; c_1, \dots, c_n) \quad (10)$$

involving an n -vector $c = (c_1, \dots, c_n)$ as a parameter be admissible for any c . If $J[u_n(\cdot; c)] = F(c)$, obtained by substituting u_n into J , is minimized at $c = c^0$, then $u_n(\cdot; c^0)$ is considered as a function that approximates u_0 most precisely within the family $u_n(\cdot; c)$. This vector c^0 is obtained, in general, by solving the simulta-

neous equations

$$\partial J[u_n(\cdot; c)] / \partial c_j = 0, \quad j = 1, \dots, n. \quad (11)$$

The function u_n appearing in (10) is often taken to be a so-called linear admissible function.

For instance, in the example above, let $\{\varphi_k\}_{k=1}^{\infty}$ be a system of independent functions complete in \tilde{A}_j , and set

$$u_n = c_1 \varphi_1 + \dots + c_n \varphi_n. \quad (12)$$

The method that constructs a minimizing sequence u_1, u_2, \dots by determining the value of the c_k in (12) by (11) is called **Ritz's method**, and φ_k is called a coordinate function in this method. As for the rate of convergence in approximation by Ritz's method, as well as for the estimation of errors, there are several results by the Soviet school in addition to those of E. Trefftz [3, 5]. (Other methods of constructing minimizing sequences are stated in detail in [3]; concerning a connection with [†]Galerkin's method → 304 Numerical Solution of Partial Differential Equations.)

Ritz's method applied to eigenvalue problems is called the **Rayleigh-Ritz method**. Since a stationary value of the Rayleigh quotient is itself an eigenvalue in this case, the precision of approximation is far better for eigenvalues than for eigenfunctions, so that it is a convenient method for the approximate computation of eigenvalues.

References

- [1] R. Courant and D. Hilbert, Methods of mathematical physics, Interscience, I, 1953; II, 1962.
- [2] P. Funk, Variationsrechnung und ihre Anwendungen in Physik und Technik, Springer, 1962.
- [3] L. V. Kantorovich and V. I. Krylov, Approximation methods of higher analysis (in Russian), Gostekhizdat, 1952.
- [4] S. G. Mikhlin, Variational methods in mathematical physics, Macmillan, 1964. (Original in Russian, 1957.)
- [5] S. G. Mikhlin, The problem of the minimum of a quadratic functional, Holden-Day, 1965. (Original in Russian, 1952.)
- [6] S. L. Sobolev, Applications of functional analysis in mathematical physics, Amer. Math. Soc. Transl. of Math. Monographs, 1963. (Original in Russian, 1950.)
- [7] E. Goursat, Cours d'analyses mathématique, III, Gauthier-Villars, 1927.
- [8] C. Carathéodory, Calculus of variations and partial differential equations of the first order, Holden-Day, 1967.
- [9] O. Bolza, Vorlesungen über Variationsrechnung, Teubner, 1909.

[10] I. M. Gel'fand and S. V. Fomin, *Calculus of variations*, Prentice-Hall, 1963. (Original in Russian, 1961.)

[11] C. B. Morrey, *Multiple integrals in the calculus of variation*, Springer, 1966.

47 (XXI.1 5) Cantor, Georg

George Cantor (March 3, 1845–January 6, 1918), the founder of set theory, was born in St. Petersburg into a Jewish merchant family that settled in Germany in 1856. He studied mathematics, physics, and philosophy in Zurich and at the University of Berlin. After receiving his degree in 1867 in Berlin, he became a lecturer at the University of Halle and served as a professor at that university from 1879 to 1905. In 1884, under the strain of opposition to his ideas and his efforts to prove the \dagger continuum hypothesis, he suffered the first of many attacks of depression which continued to hospitalize him from time to time until his death.

The thesis he wrote for his degree concerned the theory of numbers; however, he arrived at set theory from his research concerning the uniqueness of \dagger trigonometric series. In 1874, he introduced for the first time the concept of \dagger Cardinal numbers, with which he proved that there were “more” \dagger transcendental numbers than \dagger algebraic numbers. This result caused a sensation in the mathematical world and became the subject of a great deal of controversy. Cantor was troubled by the opposition of L. \dagger Kronecker, but he was supported by J. W. R. \dagger Dedekind and G. Mittag-Leffler. In his note on point-set theory, he wrote, in connection regard with his concept of infinity, “The essence of mathematics lies in its freedom!” In addition to his work on cardinal numbers, he laid the basis for the concepts of \dagger order types, \dagger transfinite ordinals, and the theory of real numbers by means of \dagger fundamental sequences. He also studied general point sets in Euclidean space and defined the concepts of \dagger accumulation point, \dagger closed set, and \dagger open set. He was a pioneer of the point-set theory that led to the development of general \dagger topology.

References

[1] G. Cantor, *Gesammelte Abhandlungen*, E. Zermelo and A. Fraenkel (eds.), Springer, 1932 (Olms, 1962).

[2] A. Schoenflies, *Die Krisis in Cantors mathematischem Schaffen*, *Acta Math.*, 50 (1927), 1-23.

48(X.31) Capacity

A. General Remarks

The electric capacity of a conductor in the 3-dimensional Euclidean space \mathbf{R}^3 is defined as the ratio of a given positive charge on the conductor to the value of the potential on the surface. This definition of capacity is independent of the given charge. The capacity of a set as a mathematical notion was defined first by N. Wiener (1924) and was developed by O. Frostman, C. J. de La Vallée Poussin, and several other French mathematicians in connection with \dagger potential theory.

B. Energy

Let Ω be a \dagger locally compact Hausdorff space and $\Phi(x, y)$ be a \dagger lower semicontinuous function on $\Omega \times \Omega$ such that $-\infty < \Phi \leq \infty$. A measure μ will mean a nonnegative \dagger Radon measure with compact \dagger support S_μ . Denote by $\Phi(x, \mu)$ the \dagger potential $\int \Phi(x, y) d\mu(y)$ of a measure μ with kernel Φ and by (μ, μ) the \dagger energy $\int \int \Phi d\mu d\mu$ of μ . Let X be a set in Ω , and denote by \mathcal{U}_X the class of normalized measures μ (i.e., of measures μ satisfying $\mu(\Omega) = 1$) with $S_\mu \subset X$. Let K be a nonempty compact set in Ω . Set $W(K) = \inf(\mu, \mu)$ for $\mu \in \mathcal{U}_K$, and $W(\emptyset) = \infty$ for the empty set \emptyset . For $\Phi(x, y) = 1/|x - y|$ in $\Omega = \mathbf{R}^3$, the general solution $u(x)$ of the \dagger Dirichlet problem (\dagger exterior problem) for the boundary function 1 in the unbounded component of $\mathbf{R}^3 - K$ is equal to the potential of an \dagger equilibrium mass-distribution. Therefore, if S is a smooth surface surrounding K and normals are drawn outward to S , the integral $-(1/4\pi) \cdot \int_S (\partial u / \partial n) d\sigma$ of the normal derivative is equal to $1/W(K)$. This is the **capacity** of K defined by N. Wiener (*J. Math. Phys.*, MIT, 3 (1924)) when K is a closed region. Vallée-Poussin (*Ann. Inst. H. Poincaré*, 2 (1932)) called the supremum of $\mu(\mathbf{R}^3)$ the Newtonian capacity of a bounded \dagger Borel set E , where μ runs through the class of measures μ with $S_\mu \subset E$ whose \dagger Newtonian potentials are not greater than 1 in \mathbf{R}^3 . If E is compact, the Newtonian capacity coincides with Wiener's capacity. For the \dagger logarithmic potential in \mathbf{R}^2 , $e^{-W(K)}$ is called the **logarithmic capacity**. When \dagger Green's function $g(z, \infty)$ with the pole at the point at infinity exists in the unbounded component of $\mathbf{R}^2 - K$, $\lim_{z \rightarrow \infty} (g(z, \infty) - \log |z|)$ is called **Robin's constant** and can be shown to be equal to $W(K)$. (For the relation between Robin's constant and \dagger reduced extremal distance \rightarrow 143 Extremal Length.) In the case of a general

kernel it is difficult to define capacity as above by means of $W(K)$, and hence the value of $W(K)$ itself instead of the capacity of K is often used. When $W(K) = \infty$, we can say that K is of capacity zero. The minimum value of the † Gauss integral $(\mu, \mu) = 2 \iint d\mu$ is a generalization of $W(K)$, where $\mu \in \mathcal{U}_K$ and f is an upper semicontinuous function bounded above on K .

C. Minimax Value

Suppose that we are given the kernel Φ as above. For a set $X \subset \Omega$ and a measure μ , set $U(\mu; X) = \sup_{x \in X} \Phi(x, \mu)$ and $V(\mu; X) = \inf_{x \in X} \Phi(x, \mu)$. Next, for $Y \subset \Omega$, set $U(Y) = \inf_{\mu \in \mathcal{U}_Y} U(\mu; X)$, $V(Y) = \sup_{\mu \in \mathcal{U}_Y} V(\mu; X)$, $U(X, Y) = \inf_{\mu \in \mathcal{U}_Y} U(\mu; X)$, and $V(X, Y) = \sup_{\mu \in \mathcal{U}_Y} V(\mu; X)$, where $\mu \in \mathcal{U}_Y$. If $\check{\Phi}(x, y) = \Phi(x, y)$ is taken as a kernel instead of $\Phi(x, y)$, then the notations $u(K)$, $\check{U}(\mu; X)$, $\check{V}(\mu; X)$, are used correspondingly. For any compact set K the following relations hold:

$$W(K) = \check{W}(K) \leq CI(K) = \check{u}(K)$$

$$\leq \left\{ \begin{array}{l} U(K, K) = \check{V}(K, K) \\ \check{U}(K, K) = V(K, K) \end{array} \right\} \leq \left\{ \begin{array}{l} U(\Omega, K) = \check{V}(K, \Omega) \\ \check{U}(\Omega, K) = V(K, \Omega) \\ V(K) = \check{V}(K) \end{array} \right\}$$

Examples show that all the inequalities can be strict. The † minimax theorem in the † theory of games plays an important role in the proof of these inequalities [7]. Even if the kernel is symmetric, the inequalities can be strict except for the equality $W(K) = U(K)$. When the kernel is positive, we can define the quantities which correspond to $U(Y)$, $V(Y)$, $U(X, Y)$, $V(X, Y)$ by considering the class of μ with $S_{\mu} \subset Y$ and $(\mu, \mu) = 1$ instead of \mathcal{U}_Y .

D. Transfinite Diameter

As $k \rightarrow \infty$,

$$D_k(K) = k^{-1}(k-1)^{-1} \inf_{x_1, \dots, x_k \in K} \sum_{i \neq j} \Phi(x_i, x_j)$$

decreases and the limit $D(K)$ is equal to $W(K)$. For the logarithmic kernel in \mathbf{R}^2 , M. Fekete defined $D(K)$ and called $e^{-D(K)}$ the **transfinite diameter** of K (1923). F. Leja and his school in Poland studied relations between transfinite diameter and † conformal mapping. Next, set

$$kR_k(X, Y) = \sup_{x_1, \dots, x_k \in X} \inf_{y \in Y} \sum_{i=1}^k \Phi(x_i, y).$$

Then $R(X, Y) = \lim_{k \rightarrow \infty} R_k(X, Y)$ exists as $k \rightarrow \infty$, and we have $R(K, Y) = V(K, Y)$.

Fekete introduced $R(K) = R(K, K)$ in \mathbf{R}^2 (1923). G. Polya and G. Szegő computed $D(K)$ and $R(K)$ for special K and cc-kernel $r^{-\alpha}$ ($\alpha \geq 0$) in \mathbf{R}^2 and \mathbf{R}^3 [12]. The equality $D(K) = R(K)$ holds for the logarithmic kernel in \mathbf{R}^2 and the

Newtonian kernel in \mathbf{R}^3 . The maximum of the absolute value on K of † Chebyshev's polynomial (\rightarrow 336 Polynomial Approximation) of order k with respect to K in \mathbf{R}^2 is equal to $\exp(-kR_k(K))$.

E. Evans's Theorem

In order that K be of Newtonian capacity zero, it is necessary and sufficient that there exist a measure μ on K such that the Newtonian potential of μ is equal to ∞ at every point of K . This result was proved by G. C. Evans and H. Selberg independently (1935) and is called **Evans's theorem** (or the **Evans-Selberg theorem**). The corresponding theorem in \mathbf{R}^2 is often applied in the theory of functions [15, 16]. A similar potential exists in case of a general kernel if and only if $R(K, K) = \infty$.

F. Nonadditivity of Capacity

Many kinds of capacity satisfy the inequality $\text{cap}(\bigcup_n X_n) \leq \sum_n \text{cap } X_n$, where a capacity is denoted by cap . Even the Newtonian capacity C is not necessarily additive, but it satisfies $C(K_1 \cup K_2) + C(K_1 \cap K_2) \leq C(K_1) + C(K_2)$ (G. Choquet [2]). Choquet [3] proved that X can be divided into mutually disjoint sets X , and X_2 such that $C_i(X) = C_i(X_1) = C_i(X_2)$, where $C_i(X)$ is the **Newtonian inner** (or **interior**) **capacity** defined to be $\sup_{K \subset X} C(K)$ if $X \neq \emptyset$ and 0 if $X = \emptyset$.

G. Relation to Hausdorff Measure

There are many studies of relations between capacity and † Hausdorff measure [1]. Frostman [6] introduced the notion of **capacity dimension** and observed that it coincides with the Hausdorff dimension. The capacity of product sets has been evaluated from above and below [11]. For compact sets $K \subset \mathbf{R}^n$, $K' \subset \mathbf{R}^m$, their dimensions α , β , and the dimension γ of $K \times K'$, we have the relation $\alpha + \beta \leq \gamma \leq \min(m + \alpha, n + \beta)$, where the equalities are attained by general † Cantor sets. There are also works on the evaluation of capacities of general Cantor sets [10, 15]. If K is a continuum of logarithmic capacity 1 in a plane, then its diameter d satisfies $2 \leq d \leq 4$, and its area A satisfies $A < n$ [S]. Consider the sum $K = \{z_1 + \dots + z_n | z_k \in K_k, 1 \leq k \leq n\}$ of continua K_1, \dots, K_n in a plane. The logarithmic capacity of K is strictly greater than the sum of the logarithmic capacities of K_1, \dots, K_n except when all K_k are convex and mutually similar [14]. By various † symmetrizations the logarithmic capacity decreases in general (\rightarrow 228 Isoperimetric Problems; also [133]).

H. Capacitability

The **Newtonian outer** (or **exterior**) capacity $C_*(X)$ is defined by $\inf C_i(G)$, where G ranges over an open set containing X . The inequality $C_i(X) \leq C_*(X)$ holds, in general. When the equality holds, X is called **capacitable**. Choquet (1955) [2] proved that all \aleph_1 -analytic sets and hence \aleph_1 -Borel sets are capacitable but there exists an analytic set whose complement is not capacitable. He himself generalized his result on capacitability in the following way [4]: Let Ω be an abstract space, φ a nondecreasing function defined on the family of all subsets of Ω , and \mathcal{H} some family of subsets of Ω that is closed under the formation of finite unions and countable intersections. Assume that $\varphi(H_n) \downarrow \varphi(H)$ as H_n in \mathcal{H} decreases to H and that $\varphi(X_n) \uparrow \varphi(X)$ as $X_n \uparrow X$. When $\varphi(X)$ is equal to $\sup\{\varphi(H) \mid H \in \mathcal{H}, H \subset X\}$, X is called (φ, \mathcal{H}) -capacitable. Choquet defined \mathcal{H} -Suslin sets and showed that they are (φ, \mathcal{H}) -capacitable. M. Kishi [9], Choquet [5], and B. Fuglede [7] investigated capacitability with respect to several kinds of capacity more general than Newtonian capacity. We can discuss capacitability with respect to quantities defined in connection with the \aleph_1 -Gauss variational problem.

References

- [1] L. Carleson, Selected problems on exceptional sets, Van Nostrand, 1967.
- [2] G. Choquet, Theory of capacities, Ann. Inst. Fourier, 5 (1955), 131-295.
- [3] G. Choquet, Potentiels sur un ensemble de capacité nulle, Suites de potentiels, C. R. Acad. Sci. Paris, 244 (1957), 1707-1710.
- [4] G. Choquet, Forme abstraite du théorème de capacitabilité, Ann. Inst. Fourier, 9 (1959), 83-89.
- [5] G. Choquet, Diamètre transfini et comparaison de diverses capacités, Sémin. Théorie du Potentiel, 3 (1958-1959), no. 4.
- [6] O. Frostman, Potentiel d'équilibre et capacité des ensembles avec quelques applications à la théorie des fonctions, Medd. Lunds Univ. Mat. Sem., 3 (1935).
- [7] B. Fuglede, Une application du théorème du minimax à la théorie du potentiel, Colloque Internat. sur la Théorie du Potentiel, Paris, Orsay, 1964, C. N. R. S. exposé no. 8.
- [8] G. M. Goluzin, Geometrische Funktionentheorie, Deutscher Verlag der Wiss., 1957.
- [9] M. Kishi, Capacitability of analytic sets, Nagoya Math. J., 16 (1960), 91-109.
- [10] M. Ohtsuka, Capacité d'ensembles de Cantor généralisés, Nagoya Math. J., 11 (1957), 151-160.

- [11] M. Ohtsuka, Capacité des ensembles produits, Nagoya Math. J., 12 (1957), 95-130.
- [12] G. Pólya and G. Szegő, Über den transfiniten Durchmesser (Kapazitätskonstante) von ebenen und räumlichen Punktmengen, J. Reine Angew. Math., 165 (1931), 4-49.
- [13] G. Pólya and G. Szegő, Isoperimetric inequalities in mathematical physics, Princeton Univ. Press, 1951.
- [14] C. Pommerenke, Zwei Bemerkungen zur Kapazität ebener Kontinuen, Arch. Math., 12 (1961), 122-128.
- [15] M. Tsuji, Potential theory in modern function theory, Maruzen, 1959.
- [16] L. Sario and K. Oikawa, Capacity functions, Springer, 1969.

49 (11.6)
Cardinal Numbers

A. Definition

The general concept of cardinal number is an extension of that of natural number (Cantor [1]). When there exists a \aleph_1 -one-to-one correspondence whose \aleph_1 -domain is a set A and whose \aleph_1 -range is a set B , this set B is said to be **equipotent** (or **equipollent**) to A , and this relation is denoted by $A \sim B$. The relation \sim is an \aleph_1 -equivalence relation, and each equivalence class under this relation is said to be a **cardinal number**. The class of all sets equipotent to a set A is denoted by A (or $|A|$) and is said to be the **cardinal number** (power, cardinality, or **potency**) of the set A . When A is a finite set, A is said to be **finite**, and when A is an infinite set, A is said to be **infinite** (or **transfinite**). When the cardinal number of a set A is m , A is also said to consist of m members (or m elements). In this sense, 0 and the natural numbers are considered to express finite cardinal numbers. For example, $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, etc. Examples of infinite cardinal numbers: A set A which is equipotent to the set N of all natural numbers is said to be **countably infinite**, and the cardinal number of the set N is denoted by \aleph_0 . A set A which is finite or countably infinite is said to be **countable**. The cardinal number of the set of all real numbers is denoted by c and is called the **cardinal number of the continuum**. Moreover, the cardinal number of the set of all real-valued functions whose domain is the interval $[0, 1]$ is denoted by \mathfrak{f} . These three cardinal numbers are known to be distinct. Henceforth in this article, lower-case German letters denote cardinal numbers. For a definition of cardinal numbers using the concept of ordinal numbers \rightarrow 312 Ordinal Numbers.

B. Ordering of Cardinal Numbers

$m \geq n$ or $n \leq m$ will mean that there exist sets A and B such that $m = \bar{A}$, $n = \bar{B}$, and $A \supset B$. $A \neq B$ does not necessarily imply $m \neq n$. For example, the cardinal number of the set of all positive even numbers is also \aleph . $m \geq n$ and $n \geq m$ imply $m = n$ (**Bernshtein's theorem**). Since the \dagger reflexive and \dagger transitive laws for the relation \leq between cardinal numbers are obvious, the relation is an \dagger ordering relation. The \dagger well-ordering theorem implies that \geq is a \dagger total ordering (**comparability theorem for cardinal numbers**). $m > n$ means that $m \geq n$ and $m \neq n$. When $B \leq m$, B is said to be at **most** m .

C. Sum, Product, and Power of Cardinal Numbers

For cardinal numbers m and n , choose sets A and B so that $m = \bar{A}$, $n = \bar{B}$, and $A \cap B = \emptyset$, and put $s = \overline{A \cup B}$. Then s is uniquely determined by m and n . The s is said to be the **sum** of the cardinal numbers m and n and is denoted by $m + n$. If the sets A, B are chosen as described above, the cardinal numbers of the \dagger Cartesian product $A \times B$ and of the set of functions A^B are called the **product** of m and n and the n th **power** of m , denoted by mn and m^n , respectively. These operations are also determined by m and n . For these three operations, the following laws are valid: **commutative laws** $m + n = n + m$, $mn = nm$; **associative laws** $(m + n) + p = m + (n + p)$, $(mn)p = m(np)$; **distributive law** $p(m + n) = pm + pn$; **exponential laws** $m^{n+p} = m^n m^p$, $m^{np} = (m^n)^p$, $(mn)^p = m^p n^p$. In particular, if $\bar{A} = m$, then 2^m is the cardinal number of the \dagger power set $\mathfrak{P}(A)$ of A .

Addition and multiplication of more than two cardinal numbers can be defined as follows. Let Λ be any set, and suppose that to any element λ of Λ there corresponds a unique cardinal number m_λ . Let M_λ be a set such that $\bar{M}_\lambda = m_\lambda$, and $M_\lambda \cap M_{\lambda'} = \emptyset$ for $\lambda \neq \lambda'$. Then the cardinal number of the \dagger disjoint union $\sum_\lambda M_\lambda$ is said to be the **sum** of all m_λ and is denoted by $\sum_\lambda m_\lambda$. The cardinal number of the Cartesian product $\prod_\lambda M_\lambda$ is said to be the **product** of all m_λ ($\lambda \in \Lambda$) and is denoted by $\prod_\lambda m_\lambda$. The axiom of choice can be stated as follows: If $m_\lambda \neq 0$ for all $\lambda \in \Lambda$, then $\prod_\lambda m_\lambda \neq 0$.

D. The Continuum Hypothesis

For \aleph , \mathfrak{c} , and \mathfrak{f} defined as before, $\mathfrak{f} = 2^{\mathfrak{c}} > \mathfrak{c} = 2^{\aleph}$. In general, $2^m > m$ holds for any cardinal

number m (Cantor). The hypothesis which asserts that for any m , there does not exist an n such that $2^m > n > m$ is called the **generalized continuum hypothesis**. In particular, this hypothesis restricted to the case where $m = \aleph$ is called the **continuum hypothesis**. After Cantor stated this hypothesis (*J. Reine Angew. Math.*, 84 (1878)), it remained an open question for many years. In particular, Cantor himself repeatedly tried to prove it, and W. Sierpiński pursued various related hypotheses. Finally, the continuum hypothesis and the generalized continuum hypothesis were proved to be independent of the axioms of set theory by K. Gödel (1940) [3] and P. J. Cohen (1963) [4] (\rightarrow 33 Axiomatic Set Theory).

E. Cardinality of Ordinal Numbers

Lower-case Greek letters will stand here for \dagger ordinal numbers. The cardinal number of $\{\xi \mid \xi < \alpha\}$ will be denoted by $\bar{\alpha}$, which is called the **cardinality of the ordinal number α** or the **cardinal number corresponding to α** . When a cardinal number m corresponds to some ordinal number, the minimum among ordinal numbers α with $\bar{\alpha} = m$ is called the **initial ordinal number** corresponding to m . An initial ordinal number corresponding to an infinite cardinal number is called a **transfinite initial ordinal number**. There exists a unique correspondence $\beta \rightarrow \omega_\beta$ from the class of ordinal numbers onto the class of all transfinite initial ordinal numbers such that $\beta > \gamma$ implies $\omega_\beta > \omega_\gamma$. In particular, $\omega_0 = \omega$, and an ordinal number ξ such that $\xi < \omega_1$ is called a **countable ordinal number**. ω_β is called the β th transfinite initial ordinal number. The cardinality of ω_β is denoted by \aleph_β (\aleph is the Hebrew letter **aleph**). In particular, \aleph is denoted by \aleph_0 (**aleph zero**). We have $\aleph_\beta \geq \aleph_\gamma$ if and only if $\beta \geq \gamma$, and in this case $\aleph_\beta + \aleph_\gamma = \aleph_\beta$, $\aleph_\beta \aleph_\gamma = \aleph_\beta$. The axiom of choice implies that every infinite cardinal number is an \aleph_β . Hence, in this case, the continuum hypothesis can be formulated as $2^{\aleph_0} = \aleph_1$, and the generalized continuum hypothesis can be formulated as $2^{\aleph_\beta} = \aleph_{\beta+1}$ for every ordinal number β .

F. Finiteness and Infiniteness

Dedekind [5] defined a set A to be **infinite** if A is equipotent to a proper subset of itself, and to be **finite** otherwise. It is also possible to define finiteness and infiniteness of sets as follows: A set A is finite if there exists a \dagger well-ordering of A such that its \dagger dual ordering is also a well-ordering, and A is infinite other-

wise. If a set is finite in the latter sense, then it is also finite in the sense of Dedekind. Under the axiom of choice, these two definitions can be shown to be equivalent.

References

- [1] G. Cantor, Beiträge zur Begründung der transfiniten Mengenlehre I, Math. Ann., 46 (1895), 471–512. Gesammelte Abhandlungen, Springer, 1932; English translation, Contributions to the founding of the theory of transfinite numbers, Open Court, 1915.
 - [2] W. Sierpiński, Hypothèse du continu, Warsaw, 1934.
 - [3] K. Gödel, The consistency of the continuum hypothesis, Ann. Math. Studies, Princeton Univ. Press, 1940.
 - [4] P. J. Cohen, Set theory and the continuum hypothesis, Benjamin, 1966.
 - [5] R. Dedekind, Was sind und was sollen die Zahlen? Vieweg, 1888. Gesammelte mathematische Werke 3, Vieweg, 1932; English translation, Essays on the theory of numbers, Open Court, 1901.
- Also → references to 381 Sets.

50 (XXI.16) Cartan, Elie

Elie Cartan (April 9, 1869–May 6, 1951) was born at Dolomieu in the French province of Isère. He entered the Ecole Normale Supérieure in Paris in 1888 and graduated in 1891, having at the same time qualified in the agrégé examination. Beginning his research immediately, he completed his thesis on the structure of continuous transformation groups [2] in 1894 at the age of 25.

Cartan was a professor first at the University of Montpellier, later at the University of Lyon, then the University of Nancy, and finally in 1912 at the University of Paris. He freely used the †moving coordinate system introduced by J. G. Darboux, and contributed to many areas, such as the theory of †Lie groups, the theory of †Pfaffian forms, the theory of †invariant integrals, †topology, †differential geometry (especially the geometry of †connections), and theoretical physics. His doctoral thesis is still an object of interest among young researchers today, and the concept of connection introduced by him is fundamental in the field of differential geometry. Henri Cartan (1904–) is his eldest son.

References

- [1] E. Cartan, Oeuvres complètes I–III, Gauthier-Villars, 1952–1955.
- [2] E. Cartan, Sur la structure des groupes de transformations finis et continus, Thesis, 1894. (Oeuvres complètes, pt. I, vol. 1, 137–287.)
- [3] S. S. Chern and C. Chevalley, Elie Cartan and his mathematical work, Bull. Amer. Math. Soc., 58 (1952), 217–250.

51 (IX.24) Catastrophe Theory

A. General Remarks

Catastrophe theory was originally proposed by R. Thom [1, 2] in late 1960. This theory provides certain mathematical models for the evolution of forms in nature, in particular in biology and the natural languages. In recent years, E. C. Zeeman and others have developed and applied the theory to various fields, including the physical sciences, medicine, economics, and sociology; a detailed bibliography may be found in [3, 9, 10].

B. Static Models

The basic concept of the theory is the **static model**. This is a family of potential functions $f_u: X \rightarrow \mathbf{R}$, where X is a subset of \mathbf{R}^n , containing a neighborhood of the origin, and the parameter u lies in a neighborhood U of the origin in \mathbf{R}^r .

We regard \mathbf{R}^n as the **internal space** or **state space**, which is parametrized by the various variables that are relevant to the process under study, and \mathbf{R}^r is the **external space** or **control space**, which can be either the physical space-time continuum in which the process under consideration takes place or a space of control parameters that govern the process. As a rule, we assume that the dimension $r \leq 4$, although the dimension n can be arbitrary large.

The static model is local in nature, and any one of the local minima of f_u , called a **local regime** at $u \in U$, is a candidate for the state of the model corresponding to the control point u .

Mathematically, a static model is a germ of C^∞ -functions $f: \mathbf{R}^n \times \mathbf{R}^r \rightarrow \mathbf{R}$ at 0 that is an unfolding (r -dimensional extension family) of a germ of C^∞ -functions $\eta = f|_{\mathbf{R}^n \times \{0\}}: \mathbf{R}^n \rightarrow \mathbf{R}$ at 0; further details are given in Section D.

C. Classification of Singularities

Let $\mathcal{E}(n, m)$ be the vector space of \dagger germs of C^∞ -functions $f: \mathbf{R}^n \rightarrow \mathbf{R}^m$ at 0, and let $\mathcal{B}(n) \subset \mathcal{E}(n, n)$ be the subset of invertible germs $\mathbf{R}^n \rightarrow \mathbf{R}^n$ mapping 0 to 0.

The germs $\eta, \xi \in \mathcal{E}(n, m)$ are called **right equivalent** if there exists an $h \in \mathcal{B}(n)$ such that $\eta \circ h = \xi$. Let $\mathcal{E}(n) = \mathcal{E}(n, 1)$. Then it follows in the usual manner that $\mathcal{E}(n)$ is a local algebra with the unique maximal ideal $\mathcal{M} = \mathcal{M}(n) = \{\eta \in \mathcal{E}(n) \mid \eta(0) = 0\}$. A germ $\eta \in \mathcal{E}(n)$ such that $\eta(0) = D\eta(0) = 0$ is called a **singularity**. For any singularity η , we define the **codimension** of η by $\text{codim } \eta = \dim_{\mathbf{R}}(\mathcal{M} / \langle \partial \eta / \partial x_i \rangle_{\mathcal{E}(n)})$,

where $\langle \partial \eta / \partial x_i \rangle_{\mathcal{E}(n)}$ is the ideal of $\mathcal{E}(n)$ generated by $\partial \eta / \partial x_1, \dots, \partial \eta / \partial x_n$ and $x = (x_1, \dots, x_n)$ denotes the coordinate system of \mathbf{R}^n .

The following result was proved by J. N. Mather [4]: Up to the addition of a nondegenerate quadratic form in other variables, and up to multiplication by ± 1 , a singularity of codimension ≤ 4 and ≥ 1 is right equivalent to one of the η appearing in Thom's list of the seven elementary catastrophes below.

D. Unfoldings

Let η be a singularity. An **r -unfolding** of η is a germ $f \in \mathcal{M}(n+r)$ such that $f|_{\mathbf{R}^n \times \{0\}} = \eta$; this unfolding is denoted by (r, f) . Let (r, f) and (s, g) be unfoldings of η . A **morphism** $(\varphi, \Phi, \varepsilon): (r, f) \rightarrow (s, g)$ consists of (i) a germ $\varphi \in \mathcal{E}(n+r, n+s)$ such that $\varphi|_{\mathbf{R}^n \times \{0\}} = \text{identity}$, (ii) a germ $\Phi \in \mathcal{E}(r, s)$ such that $\pi_s \circ \varphi = \Phi \circ \pi_r$, (iii) a germ $\varepsilon \in \mathcal{M}(r)$ such that $f = g \circ \varphi + \varepsilon \circ \pi_r$, where $\pi_r: \mathbf{R}^n \times \mathbf{R}^r \rightarrow \mathbf{R}^r$ is the projection. In this case we say that the unfolding (r, f) is **induced** by $(\varphi, \Phi, \varepsilon)$ from (s, g) . A morphism is an **isomorphism** if φ and Φ are diffeomorphic germs.

The **addition** of unfoldings (r, f) and (s, g) of η is defined by $(r, f) + (s, g) = (r+s, f+g-\eta)$, where the last term on the right is given by $(f+g-\eta)(x, u, v) = f(x, u) + g(x, v) - \eta(x)$.

Thus if the **constant unfolding** (s, η) of η is

defined by $\eta(x, u) = \eta(x)$, we have $(r, f) + (s, \eta) = (r+s, f)$.

An unfolding (r, f) of η is said to be **versal** if any unfolding of η is induced by (r, f) and a suitable morphism. A versal unfolding (r, f) with minimal r is said to be **universal**. The following facts have also been proved by Mather [4]:

A singularity $\eta \in \mathcal{M}(n)$ has a versal unfolding if and only if $\text{codim } \eta$ is finite. Any two r -versal unfoldings of η are isomorphic. Every versal unfolding is isomorphic to $(r, f) + \text{constant}$, where $r = \text{codim } \eta$ and (r, f) is a universal unfolding defined as follows: If $\{b_1, \dots, b_r\} \subset \mathcal{M}(n)$ is a system of representatives for a basis of $\mathcal{M}(n) / \langle \partial \eta / \partial x_i \rangle_{\mathcal{E}(n)}$, then the unfolding f of η is defined by $f(x, u) = \eta(x) + b_1(x)u_1 + \dots + b_r(x)u_r$.

E. The Seven Elementary Catastrophes

Let f and g be germs in $\mathcal{E}(n+r)$. We say that f and g are **equivalent** as r -unfoldings if there exist $h \in \mathcal{B}(r)$, a family of $H_u \in \mathcal{B}(n)$, where $u \in U \subset \mathbf{R}^r$, and an $\varepsilon \in \mathcal{M}(r)$ such that $f(x, y) = g(H_u(x), h(u)) + \varepsilon(u)$.

We say that a static model (r, f) is **stable** if any small perturbation (r, g) of (r, f) in $\mathcal{E}(n+r)$ (with the Whitney C^∞ -topology) is equivalent to (r, f) . The following is the main result observed by Thom and proved by Mather and others (for references see the bibliography in [3]). Suppose that $r \leq 4$. Then the set of stable static models (r, f) is an open dense subset of $\mathcal{E}(n+r)$, and up to the addition of a nondegenerate quadratic form and multiplication by ± 1 , any stable static model (r, f) is equivalent to one of the models with the standard potentials F , which are the universal unfoldings of singularities η in Table 1 (x, y denote internal variables, and u, v, w, t denote external variables).

The static models with these standard potentials are known as the **elementary catastrophes** and can serve as qualitative models of various natural processes.

Table 1. Thom's List of the Seven Elementary Catastrophes

r	Singularity η	Standard Potential F	Name
1	x^3	$x^3 + ux$	Fold
2	x^4	$x^4 + ux^2 + vx$	Cusp (Riemann-Hugo-Niot)
3	x^5	$x^5 + ux^3 + vx^2 + wx$	Swallowtail (Dovetail)
3	$x^3 + y^3$	$x^3 + y^3 + uxy + vx + wy$	Hyperbolic umbilic
3	$x^3 - xy^2$	$x^3 - xy^2 + u(x^2 + y^2) + vx + wy$	Elliptic umbilic
4	x^6	$x^6 + tx^4 + ux^3 + vx^2 + wx$	Butterfly
4	$x^2y + y^4$	$x^2y + y^4 + ux^2 + vy^2 + wx + ty$	Parabolic umbilic

F. Catastrophe Sets

A **process** for the static model $f \in \mathcal{E}(n+r)$ is a subset s of $X \times U$, where X and U are neighborhoods of the origin in \mathbf{R}^n and \mathbf{R}^r , respectively. If s is a process and $u \in U$, we define $s_u = s \cap (X \times \{u\})$. We say that $u \in U$ is a **regular point** for s if there exists a neighborhood V of u in U and a homeomorphism $h: X \times V \rightarrow X \times V$ such that $\pi_r \circ h = \pi_r$ on $X \times V$ and $h(s \cap (X \times V)) = s_u \times V$. A **catastrophe point** is a non-regular point in U . The set of all catastrophe points is called the **catastrophe set**.

There are several conventions with regard to the definition of a process for the unfolding (r, f) . One of these is the **Maxwell convention**, which requires that s_u be a least local regime at $u \in U$. The catastrophe set of the model (r, f) according to the Maxwell convention consists of those points u of the control space \mathbf{R}^r where the least minimum of f_u is attained for at least two points or where this minimum is attained at a unique point but is not stable.

Another is the **perfect delay convention**, which assigns to each path τ in U a mapping (possibly discontinuous) $m_\tau: \tau \rightarrow X \times \tau$ such that $\pi_r(m_\tau(u)) = u$, $(m_\tau(u), u)$ is a local regime, and m_τ remains continuous for a maximum interval on the path τ . Consider the set

$$\Delta = \{(x, u) \in \mathbf{R}^n \times \mathbf{R}^r \mid d_x^2 f(x, u) \text{ is degenerate}\}$$

and its image $B = \pi_r(\Delta)$, called the **bifurcation set**, under the projection $\pi_r: \mathbf{R}^n \times \mathbf{R}^r \rightarrow \mathbf{R}^r$. Then the points of the bifurcation set B are important candidates for catastrophe points of the static model with respect to the perfect delay convention.

For geometrical studies of the elementary catastrophes \rightarrow [6–8]. The static models have been generalized to **metabolic models**, the structure of which largely remains an open question [2]. In this connection, the bibliographies of [3, 9, 10] appear to be fairly complete.

References

- [1] R. Thom, Topological models in biology, *Topology*, 8 (1969), 313–335.
- [2] R. Thom, *Stabilité structurelle et morphogénèse*, Benjamin, 1972; English translation, *structural stability and morphogenesis*, Benjamin, 1975.
- [3] A. Manning (ed.), *Dynamical systems—Warwick 1974*, Lecture notes in math. 468, Springer, 1975.
- [4] J. N. Mather, Right equivalence, Warwick preprint, 1969.
- [5] G. Wassermann, Stability of unfoldings, Lecture notes in math. 393, Springer, 1974.

- [6] T. Bröcker, Differentiable germs and catastrophes, London Math. Soc. Lecture Notes Series 17, Cambridge Univ. Press, 1975.
- [7] A. N. Godwin, Three-dimensional pictures for Thom's parabolic umbilic, *Publ. Math. Inst. HES*, 40 (1971), 117–138.
- [8] A. E. R. Woodcock and T. Poston, A geometrical study of the elementary catastrophes, Lecture notes in math. 373, Springer, 1974.
- [9] T. Poston and I. Stewart, *Catastrophe theory and its applications*, Pitman, 1978.
- [10] E. C. Zeeman, *Catastrophe theory, Selected papers 1972–1977*, Addison-Wesley, 1977.

52 (II.24)
Categories and Functors

A. Categories

Consider the family of all \dagger groups. Given two groups X and Y , denote the set of all \dagger homomorphisms of X to Y by $\text{Hom}(X, Y)$. If X, Y , and Z are groups and if $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ are homomorphisms, we can compose them to get a homomorphism $g \circ f: X \rightarrow Z$.

In general, suppose that we are given, as in this example, (1) a family \mathfrak{M} of **mathematical objects**, and (2) for every pair (X, Y) of objects in \mathfrak{M} , a set $\text{Hom}(X, Y)$ whose elements are called **morphisms**; and let $f \in \text{Hom}(X, Y)$ and $g \in \text{Hom}(Y, Z)$ determine a morphism $g \circ f \in \text{Hom}(X, Z)$, which is called their **composite**. A morphism $f \in \text{Hom}(X, Y)$ is also written $f: X \rightarrow Y$. Suppose further that these morphisms satisfy the following axioms: (1) if $f: X \rightarrow Y, g: Y \rightarrow Z$, and $h: Z \rightarrow W$ are morphisms, then $(h \circ g) \circ f = h \circ (g \circ f)$; (2) for each object $X \in \mathfrak{M}$ there exists a morphism $1_X: X \rightarrow X$ such that for any $f: X \rightarrow Y$ and $g: Z \rightarrow X$ we have $f \circ 1_X = f$ and $1_X \circ g = g$; (3) $\text{Hom}(X, Y)$ and $\text{Hom}(X', Y')$ are disjoint unless $X = X'$ and $Y = Y'$. Then we call the whole system (i.e., the family of objects \mathfrak{M} , the morphisms, and the composition of morphisms) a **category**. The elements in \mathfrak{M} are called the **objects** of the category.

By axioms (1) and (2), the set $\text{Hom}(X, X)$ is a semigroup (with respect to the composition of morphisms) which has 1_X as the identity element. Hence 1_X is determined uniquely by X . On the other hand, axiom (3) implies that a morphism f determines the objects X and Y such that $f \in \text{Hom}(X, Y)$. From these facts we can give an alternative definition of category using only the morphisms and their composition.

The totality of the objects (or morphisms) in a category \mathcal{C} is denoted by **Ob**(\mathcal{C}) (or **FI**(\mathcal{C});

the notation Fl comes from the French word *flèche*. The relation $x \in \text{Ob}(\mathcal{C})$ is often abbreviated to $x \in \mathcal{C}$, while $\text{Hom}(X, Y)$ is written $\text{Hom}_{\mathcal{C}}(X, Y)$ if necessary. A **subcategory** of a category \mathcal{C} is a category \mathcal{C}' with $\text{Ob}(\mathcal{C}') \subset \text{Ob}(\mathcal{C})$, such that for $X, Y \in \mathcal{C}'$ we have $\text{Hom}_{\mathcal{C}'}(X, Y) \subset \text{Hom}_{\mathcal{C}}(X, Y)$ and the composition in \mathcal{C}' is the restriction of \mathcal{C} to \mathcal{C}' . If $\text{Hom}_{\mathcal{C}'}(X, Y) = \text{Hom}_{\mathcal{C}}(X, Y)$ for all $X, Y \in \mathcal{C}'$, we say that \mathcal{C}' is a **full subcategory** of \mathcal{C} .

We define the **product category** $\mathcal{C}_1 \times \mathcal{C}_2$ of two categories in the canonical way, using the pairs of objects and the pairs of morphisms.

B. Examples of Categories

(1) Taking all sets as the objects, all mappings as the morphisms, and the composition of mappings as the composition, we obtain a category called the **category of sets**, denoted by (Sets) (or (Ens) from the French *ensemble*). For the empty set \emptyset we make the convention that $\text{Hom}(\emptyset, Y)$ contains just one element for any Y and that $\text{Hom}(Y, \emptyset)$ is empty if $Y \neq \emptyset$.

(2) As we have seen, taking all groups as the objects and the homomorphisms as the morphisms, we get the **category of groups**, written (Gr) . If we limit the objects to \dagger Abelian groups, we get the **category of Abelian groups** (Ab) as a full subcategory of (Gr) .

(3) Fix a ring R . The left R -modules and their R -linear mappings define the **category of left R -modules**, which we denote by ${}_R\mathcal{M}$. The category of right R -modules, \mathcal{M}_R , is defined similarly. When R is \dagger unitary, we usually limit the objects of ${}_R\mathcal{M}$ and \mathcal{M}_R to \dagger unitary modules. If R is commutative we can identify ${}_R\mathcal{M}$ with \mathcal{M}_R . When $R = \mathbb{Z}$ (the ring of rational integers), ${}_R\mathcal{M}$ can be identified with (Ab) . When R is a field, ${}_R\mathcal{M}$ is also called the **category of linear spaces over R** .

(4) Taking rings as objects and homomorphisms of rings as morphisms, we obtain the **category of rings**. The subcategory consisting of unitary commutative rings and unitary homomorphisms is called the **category of commutative rings** and is denoted by (Rings) .

(5) If we take \dagger differentiable manifolds as objects and differentiable mappings as morphisms, we obtain the **category of differentiable manifolds**. Similarly, for \dagger analytic manifolds and analytic mappings we obtain the **category of analytic manifolds**.

(6) Taking topological spaces as objects and continuous mappings as morphisms, we get a category called the **category of topological spaces** and denoted by (Top) . On the other hand, if we take the \dagger homotopy classes of continuous mappings as morphisms and define their composition in the natural way,

we obtain another category, which is called the **homotopy category of topological spaces**.

(7) Fix a \dagger preordered set I . Taking the elements of I as the objects and the pairs (x, y) of elements of I with $x \leq y$ as the (unique) morphism from x to y , we get a category, in which we define the composite of the morphisms (x, y) and (y, z) to be (x, z) .

In examples (1) through (6), the totality of the objects $\text{Ob}(\mathcal{C})$ is not a set, but a \dagger class (\rightarrow 381 Sets G; for the logical foundation of category theory \rightarrow [3, 9]).

C. Diagrams

If a set of arrows $\{A_{\alpha}\}$ and a set of points $\{B_{\beta}\}$ are given in such a way that each arrow A_{α} has a unique initial point and a unique endpoint, then we say that $\{A_{\alpha}, B_{\beta}\}$ is a **diagram**. (Usually, we consider the case where each point B_{β} is the initial point or the endpoint of at least one A_{α} (Fig. 1).) Let \mathcal{C} be a category and $\{A_{\alpha}, B_{\beta}\}$ a diagram. If we associate a morphism f_{α} in \mathcal{C} with each arrow A_{α} and an object $Z_{\beta} \in \mathcal{C}$ with each point B_{β} so that $f_{\alpha} \in \text{Hom}(Z_{\beta}, Z_{\gamma})$ whenever A_{α} has the initial point B_{β} and the endpoint B_{γ} , then we say that $\{f_{\alpha}, Z_{\beta}\}$ is a **diagram in the category \mathcal{C}** (Fig. 2). Suppose, furthermore, that the following condition is satisfied: For any pair of points B_{β} and B_{γ} , and for any sequence of adjacent arrows $A_{\alpha_1}, \dots, A_{\alpha_m}$ starting at B_{β} and ending at B_{γ} (i.e., the initial point of A_{α_1} is B_{β} , the endpoint of A_{α_i} is the initial point of $A_{\alpha_{i+1}}$ and the endpoint of A_{α_m} is B_{γ}), the composite $f_{\alpha_m} \circ f_{\alpha_{m-1}} \circ \dots \circ f_{\alpha_1} (\in \text{Hom}(B_{\beta}, B_{\gamma}))$ depends only on B_{β} and B_{γ} . Then the diagram in \mathcal{C} is said to be a **commutative diagram**. For example, commutativity of Fig. 2 is equivalent to $f_3 \circ f_1 = f_4 \circ f_2 = f_5$.

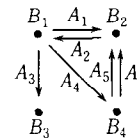


Fig. 1

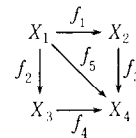


Fig. 2

D. Miscellaneous Definitions

A morphism $f: X \rightarrow Y$ in a category \mathcal{C} is called an **isomorphism** (or **equivalence**) if there exists a morphism $g: Y \rightarrow X$ such that $f \circ g = 1_Y$, $g \circ f = 1_X$. In this case, g is determined uniquely by f and is itself an isomorphism. We call g the **inverse morphism** of f . Then the inverse of g is f . An isomorphism is sometimes written $f: X \simeq Y$. Two objects X and Y are said to be **isomorphic** if there is an isomorphism $X \rightarrow Y$, and then we write $X \cong Y$. The composite of

isomorphisms is again an isomorphism. In particular, an isomorphism $X \rightarrow X$ is an invertible element of the semigroup $\text{Hom}(X, X)$, and is called an **automorphism** of X . Examples of isomorphisms are bijections in (Sets), group isomorphisms in (Gr), the R -isomorphisms in $R\mathcal{M}$, ring isomorphisms in the category of rings, \dagger diffeomorphisms in the category of differentiable manifolds, and homeomorphisms in (Top).

A morphism $f: X \rightarrow Y$ is called a **monomorphism** (or **injection**) if for any object Z and for any morphisms $u, v: Z \rightarrow X (u \neq v)$ we have $f \circ u \neq f \circ v$. Dually, $f: X \rightarrow Y$ is called an **epimorphism** (or **surjection**) if for any $u, v: Y \rightarrow Z (u \neq v)$ we have $u \circ f \neq v \circ f$. In the category of sets the monomorphisms and the epimorphisms coincide, respectively, with the injections and the surjections as mappings (\rightarrow 381 Sets). A monomorphism which is at the same time an epimorphism is called a **bijection**. An isomorphism is always a bijection, but the converse is false in some categories.

Two monomorphisms $f_1: X_1 \rightarrow X$ and $f_2: X_2 \rightarrow X$ (with the same X) are said to be equivalent if there exist $g_1: X_1 \rightarrow X_2$ and $g_2: X_2 \rightarrow X_1$ such that $f_1 = f_2 \circ g_1$ and $f_2 = f_1 \circ g_2$ (Fig. 3). An equivalence class with respect to this equivalence relation is called a **subobject** of X . Similarly, we define a **quotient object** of X as an equivalence class of epimorphisms $X \rightarrow X_i$.

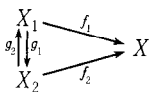


Fig. 3

An object e of a category \mathcal{C} is called a **final object** of \mathcal{C} if for every object Y of \mathcal{C} , $\text{Hom}(Y, e)$ contains one and only one element. Dually, an object e' is called an **initial object** (or **cofinal object**) if $\text{Hom}(e', Y)$ contains one and only one element for every $Y \in \mathcal{C}$. If e_1 and e_2 are final objects, then there is a unique isomorphism $e_1 \simeq e_2$, and similarly for initial objects. A set with only one element is the final object in (Sets), and a space with only one point is the final object in (Top). In the category (Gr) (resp. (Ab)), the trivial group $\{1\}$ (or $\{0\}$) is the final object and the initial object at the same time. In the category of commutative rings, the zero ring $\{0\}$ is the final object and the ring of rational integers \mathbf{Z} is the initial object.

E. Product and Coproduct

Let X_1 and X_2 be objects of a category \mathcal{C} . We say that a triple (P, p_1, p_2) consisting of an

object P and morphisms $p_i: P \rightarrow X_i (i = 1, 2)$ is the **product** (or **direct product**) of X_1 and X_2 if for any pair of morphisms $f_i: X \rightarrow X_i (i = 1, 2)$, there exists a unique morphism $f: X \rightarrow P$ with $p_i \circ f = f_i (i = 1, 2)$ (Fig. 4). If (P', p'_1, p'_2) is another product of X_1 and X_2 , then by virtue of this definition there is a unique morphism $f: P \rightarrow P'$ such that $p'_i \circ f = p_i (i = 1, 2)$, and f is an isomorphism. The product is unique in this sense. The product (or any one of the products) of X_1 and X_2 is denoted by $X_1 \times X_2$ or by $X_1 \amalg X_2$.

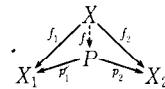


Fig. 4

The product in the categories of sets, of groups, of rings, and of topological spaces coincides with the notion of \dagger direct product in the respective systems. In a general category, the product does not always exist. Suppose the product $X \times X$ exists for an object X ; then there is a unique morphism $\Delta_X: X \rightarrow X \times X$ such that $1_X = p_1 \circ \Delta_X = p_2 \circ \Delta_X$, which is called the **diagonal morphism** of X . Let $f_i: X \rightarrow X'_i (i = 1, 2)$ be morphisms and assume that the products $(X_1 \times X_2, p_1, p_2)$, $(X'_1 \times X'_2, p'_1, p'_2)$ exist. Then there is a unique morphism $f: X_1 \times X_2 \rightarrow X'_1 \times X'_2$ satisfying $p'_i \circ f = f_i \circ p_i (i = 1, 2)$. This f is denoted by $f_1 \times f_2$. On the other hand, if $g_i: X \rightarrow X_i (i = 1, 2)$ are given, the unique morphism $g: X \rightarrow X_1 \times X_2$ with $p_i \circ g = g_i (i = 1, 2)$ is denoted by (g_1, g_2) . We have $(g_1, g_2) = (g_1 \times g_2) \circ \Delta_X$ if $X \times X$ exists.

The dual notion of product is coproduct. We say that a triple (S, j_1, j_2) of an object S and morphisms $j_i: X_i \rightarrow S (i = 1, 2)$ is the **coproduct** (or **direct sum**) of X_1 and X_2 if for any morphisms $f_i: X_i \rightarrow X (i = 1, 2)$ there exists a unique morphism $f: S \rightarrow X$ with $f \circ j_i = f_i (i = 1, 2)$ (Fig. 5). The coproduct, like the product, is uniquely determined up to canonical isomorphisms. It is denoted by $X_1 + X_2$ or by $X_1 \amalg X_2$. The coproduct in (Gr) is the \dagger free product. In (Ab), or more generally in $R\mathcal{M}$, the product of two objects can be identified with the coproduct (= direct sum) (\rightarrow 277 Modules F). The coproduct in the category of commutative rings is the \dagger tensor product over \mathbf{Z} .

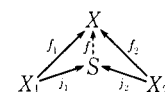


Fig. 5

Product and coproduct can also be defined for a family $\{X_i\}_{i \in I}$ of objects. Namely, the

product of $\{X_i\}_{i \in I}$ is an object P together with a family of morphisms $p_i: P \rightarrow X_i$ ($i \in I$) having the property that for any family of morphisms $f_i: X \rightarrow X_i$ ($i \in I$), there exists a unique morphism $f: X \rightarrow P$ such that $p_i \circ f = f_i$ ($i \in I$). The product is unique up to canonical isomorphisms, and similarly for the coproduct (\rightarrow Sections F, L).

F. Dual Category

In the theory of categories we often encounter the dual treatment of notions and propositions. To be precise, we may define the notion of the **dual category** \mathcal{C}° of a category \mathcal{C} as follows: The objects of \mathcal{C}° are those of \mathcal{C} , i.e., $\text{Ob}(\mathcal{C}^\circ) = \text{Ob}(\mathcal{C})$; for any objects X and Y we put $\text{Hom}_{\mathcal{C}^\circ}(X, Y) = \text{Hom}_{\mathcal{C}}(Y, X)$; if $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ in \mathcal{C}° (i.e., $f: Y \rightarrow X$ and $g: Z \rightarrow Y$ in \mathcal{C}), then the composite $g \circ f$ in \mathcal{C}° is defined to be $f \circ g$ in \mathcal{C} . It is clear that \mathcal{C}° then satisfies the axioms of a category. Quite generally, given a proposition concerning objects and morphisms we can construct another proposition by reversing the directions of the morphisms, and we call the latter the dual proposition of the former. The dual proposition of a proposition in \mathcal{C} coincides with a proposition in \mathcal{C}° . For instance, a monomorphism (epimorphism) in \mathcal{C} is an epimorphism (monomorphism) in \mathcal{C}° , and the final (initial) object in \mathcal{C} is the initial (final) object in \mathcal{C}° . The product (coproduct) in \mathcal{C} is the coproduct (product) in \mathcal{C}° . Although the notion of the dual category is defined quite formally, it is useful in describing relations between specific categories. The dual category of (Ab) , for instance, is equivalent to the category of commutative compact topological groups ('Pontryagin's duality theorem').

G. Categories over an Object

Fix a category \mathcal{C} and an object $S \in \mathcal{C}$. A pair (X, f) of an object $X \in \mathcal{C}$ and a morphism $f: X \rightarrow S$ is called an object over S or an **S -object**, and f is called its **structure morphism**. We often omit f and simply say "an S -object X " if there is no danger of misunderstanding. If (X, f) and (Y, g) are S -objects, a morphism $h: X \rightarrow Y$ such that $f = g \circ h$ is called an **S -morphism** from (X, f) to (Y, g) . The category whose objects are the S -objects and whose morphisms are the S -morphisms is called the **category of S -objects** in \mathcal{C} , and is denoted by \mathcal{C}/S . It has $(S, 1_S)$ as the final object. The product of two S -objects X and Y , taken in \mathcal{C}/S , is called the **fiber product** of X and Y over S (in \mathcal{C}), and is denoted by $X \times_S Y$ or $X \amalg_S Y$. The dual notion of the fiber product is called **fiber sum** (or **amalgamated sum**). Thus for two

morphisms $f: S \rightarrow X$ and $g: S \rightarrow Y$, the fiber product of X and Y over S in \mathcal{C}°/S is the fiber sum of X and Y (with respect to S); it is denoted by $X \amalg_S Y$.

Let \mathcal{C} be the category of commutative rings and $K \in \mathcal{C}$. Then the family of K -objects in \mathcal{C}° is precisely the family of commutative K -algebras. The fiber product $A \times_K B$ in \mathcal{C}° , i.e., $A \amalg_K B$ in \mathcal{C} , is the tensor product $A \otimes_K B$ of algebras.

H. Functors

Let \mathcal{C} and \mathcal{C}' be categories. A **covariant functor** F from \mathcal{C} to \mathcal{C}' is a rule which associates (1) with each object X in \mathcal{C} , an object $F(X)$ in \mathcal{C}' , and (2) with each morphism $f: X \rightarrow Y$ in \mathcal{C} , a morphism $F(f): F(X) \rightarrow F(Y)$ such that $F(g \circ f) = F(g) \circ F(f)$, $F(1_X) = 1_{F(X)}$. A **contravariant functor** is defined dually, by modifying this definition to $F(f): F(Y) \rightarrow F(X)$, $F(g \circ f) = F(f) \circ F(g)$. Thus a contravariant functor from \mathcal{C} to \mathcal{C}' is the same as a covariant functor from the dual category \mathcal{C}° to \mathcal{C}' (or from \mathcal{C} to \mathcal{C}'°). **Functor** is a general term for both covariant functors and contravariant functors, but some authors use the word exclusively in the sense of a covariant functor. A functor in several variables is defined to be a functor from the product category of the categories in which the variables take their values.

A covariant functor $F: \mathcal{C} \rightarrow \mathcal{C}'$ is said to be **faithful (fully faithful)** if for any $X, Y \in \mathcal{C}$, the mapping $\text{Hom}(X, Y) \rightarrow \text{Hom}(F(X), F(Y))$ induced by F is injective (bijective), and similarly for contravariant functors. A faithful covariant functor $F: \mathcal{C} \rightarrow \mathcal{C}'$ which maps distinct objects of \mathcal{C} to distinct objects of \mathcal{C}' is called an **embedding**, and in this case \mathcal{C} can be identified with a subcategory of \mathcal{C}' by F . A fully faithful covariant functor $F: \mathcal{C} \rightarrow \mathcal{C}'$ is called an **equivalence** (between the categories) if it satisfies the condition that for any object X' of \mathcal{C}' , there exists an object X of \mathcal{C} such that $F(X) \cong X'$. In this case we can consider the two categories essentially the same. A contravariant functor from \mathcal{C} to \mathcal{C}' which defines an equivalence from \mathcal{C}° to \mathcal{C}' is called an **antiequivalence**.

I. Examples of Functors

(1) Let \mathcal{C} be the category of groups (or rings). For any $X \in \mathcal{C}$ let $F(X)$ be the underlying set of X (i.e., the set obtained from X by "forgetting" its structure as a group or ring), and for any homomorphism f put $F(f) = f$. Then we get a faithful covariant functor (often called the **forgetful functor**) $F: \mathcal{C} \rightarrow (\text{Sets})$.

(2) Let \mathcal{C} be any category and fix an object X of \mathcal{C} . Then we get a covariant functor $h_X: \mathcal{C} \rightarrow (\text{Sets})$ as follows: With each $Y \in \mathcal{C}$ we associate the set $\text{Hom}(X, Y)$, and with each morphism $f: Y \rightarrow Y'$ in \mathcal{C} the mapping $f \circ$ (where $f \circ: \text{Hom}(X, Y) \rightarrow \text{Hom}(X, Y')$ is defined by $(f \circ)(g) = f \circ g$). Similarly we define a contravariant functor $h^X: \mathcal{C} \rightarrow (\text{Sets})$ by $h^X(Y) = \text{Hom}(Y, X)$ and $h^X(f) = \circ f$.

(3) Let $\rho: A \rightarrow B$ be a homomorphism of rings. With each left A -module M associate the \dagger scalar extension $\rho^*(M) = B \otimes_A M$, and with each A -homomorphism f associate the B -homomorphism $\rho^*(f) = 1_B \otimes f$. Then we get a covariant functor $\rho^*: {}_A\mathcal{M} \rightarrow {}_B\mathcal{M}$.

(4) Let R be a ring. With each left R -module M associate its dual module $M^* = \text{Hom}_R(M, R)$, and to each R -linear mapping f associate its \dagger dual mapping $f^* = \circ f$. Then we get a contravariant functor ${}_R\mathcal{M} \rightarrow {}_R\mathcal{M}$, and similarly for ${}_R\mathcal{M} \rightarrow {}_R\mathcal{M}$.

(5) For each differentiable manifold X let $F(X)$ denote the commutative ring of the differentiable functions on X , and for each differentiable mapping $f: X \rightarrow Y$ let $F(f)$ be the ring homomorphism $\circ f: F(Y) \rightarrow F(X)$. Then F is a faithful contravariant functor.

(6) Fix an Abelian group A . By associating with each topological space X the cohomology group $H(X, A)$ and with each continuous mapping $f: X \rightarrow Y$ the homomorphism $H(Y, A) \rightarrow H(X, A)$ induced by f , we obtain a contravariant functor from (Top) to (Ab) .

(7) Fix a topological space X , and let $T(X)$ be the set of the open sets in X . Then $T(X)$ is ordered by inclusion, so it is a category (\rightarrow Section B, no. 7). The contravariant functors from $T(X)$ to (Ab) are precisely the \dagger pre-sheaves of Abelian groups over X . We can use any category instead of (Ab) to define a pre-sheaf over X (\rightarrow 383 Sheaves).

J. Natural Transformations

Let \mathcal{C} and \mathcal{C}' be categories, and denote by $\text{Hom}(\mathcal{C}, \mathcal{C}')$ the collection of all covariant functors $\mathcal{C} \rightarrow \mathcal{C}'$. Let $F, G \in \text{Hom}(\mathcal{C}, \mathcal{C}')$. A **natural transformation** (or **functorial morphism**) from F to G is a function which assigns to each object X of \mathcal{C} a morphism $\varphi(X): F(X) \rightarrow G(X)$ in \mathcal{C}' such that for any morphism $f: X \rightarrow Y$ in \mathcal{C} , the equation $G(f) \circ \varphi(X) = \varphi(Y) \circ F(f)$ holds; in other words, the accompanying diagram is commutative:

$$\begin{array}{ccccc} X & & F(X) & \xrightarrow{\varphi(X)} & G(X) \\ f \downarrow & F(f) \downarrow & & & \downarrow G(f) \\ Y & & F(Y) & \xrightarrow{\varphi(Y)} & G(Y) \end{array}$$

A natural transformation between contravariant functors is defined similarly. For instance, let A and B be Abelian groups and let $H^i(\cdot, A)$ and $H^i(\cdot, B)$ be the contravariant functors of \dagger cohomology viewed as functors $(\text{Top}) \rightarrow (\text{Sets})$. Then the natural transformations between them are the \dagger cohomology operations.

Let $\varphi: F \rightarrow G$ be a natural transformation, and suppose that $\varphi(X): F(X) \rightarrow G(X)$ is an isomorphism for every $X \in \mathcal{C}$. Then the inverse transformation $G \rightarrow F$ of φ exists, and φ is called a **natural equivalence** (**functorial isomorphism** or **isomorphism**) and is written $\varphi: F \simeq G$.

Suppose that $\text{Ob}(\mathcal{C})$ is a set. Then the collection $\text{Hom}(F, G)$ of all natural transformations $F \rightarrow G$ is also a set, and hence we can consider $\text{Hom}(\mathcal{C}, \mathcal{C}')$ a category in which the objects are the covariant functors $\mathcal{C} \rightarrow \mathcal{C}'$, the morphisms are the natural transformations, and the composition of morphisms is the natural one. Then $\text{Hom}(\mathcal{C}^\circ, \mathcal{C}')$ is the category of contravariant functors from \mathcal{C} to \mathcal{C}' . In particular, the category $\text{Hom}(\mathcal{C}^\circ, (\text{Sets}))$ is sometimes denoted by \mathcal{C} .

Given a category \mathcal{C} , a covariant (resp. contravariant) functor $F: \mathcal{C} \rightarrow (\text{Sets})$, and an object $X \in \mathcal{C}$, we can define a canonical bijection $\Phi_X: \text{Hom}(h_X, F) \simeq F(X)$ (resp. $\text{Hom}(h^X, F) \simeq F(X)$) by $\Phi_X(\varphi) = \varphi(X)1_X$. (The functors h_X and h^X were defined in Section I.) The inverse mapping of Φ_X assigns to $\xi \in F(X)$ the natural transformation $\varphi: h_X \rightarrow F$ defined by $\varphi(Y)u = F(u)\xi$ ($Y \in \mathcal{C}$). In particular, if we take $F = h_Y(h^Y)$, we obtain a canonical bijection $\text{Hom}(h_X, h_Y) \simeq \text{Hom}(Y, X)$ ($\text{Hom}(h^X, h^Y) \simeq \text{Hom}(X, Y)$). It follows that there is a fully faithful contravariant (covariant) functor $\mathcal{C} \rightarrow \text{Hom}(\mathcal{C}, (\text{Sets}))$ ($\mathcal{C} \rightarrow \text{Hom}(\mathcal{C}^\circ, (\text{Sets})) = \mathcal{C}$) which associates $h_X(h^X)$ with $X \in \mathcal{C}$.

K. Adjoint Functors

Let $F: \mathcal{C} \rightarrow \mathcal{C}'$ and $F': \mathcal{C}' \rightarrow \mathcal{C}$ be covariant functors. Suppose that there is a rule which assigns to each pair of objects $M \in \mathcal{C}$ and $M' \in \mathcal{C}'$ a bijective mapping $\theta_{M, M'}: \text{Hom}_{\mathcal{C}}(M, F'(M')) \simeq \text{Hom}_{\mathcal{C}'}(F(M), M')$ such that for any pair of morphisms $N \rightarrow M$ in \mathcal{C} and $M' \rightarrow N'$ in \mathcal{C}' , the following diagram induced by the morphisms is commutative:

$$\begin{array}{ccc} \text{Hom}_{\mathcal{C}}(M, F'(M')) & \xrightarrow{\theta_{M, M'}} & \text{Hom}_{\mathcal{C}'}(F(M), M') \\ \downarrow & & \downarrow \\ \text{Hom}_{\mathcal{C}}(N, F'(N')) & \xrightarrow{\theta_{N, N'}} & \text{Hom}_{\mathcal{C}'}(F(N), N') \end{array}$$

Then we say that F is a **left adjoint functor** of F' and that F' is a **right adjoint functor** of F . We can regard $\text{Hom}_{\mathcal{C}'}(F(M), M')$ as a functor from $\mathcal{C} \times \mathcal{C}'$ (contravariant in the variable

$M \in \mathcal{C}$, covariant in the variable $M' \in \mathcal{C}'$) to the category of sets, and similarly for $\text{Hom}_{\mathcal{C}}(M, F(M'))$. This commutativity of the diagram means that these two functors are isomorphic (\rightarrow Section J).

For instance, let A and B be rings and L a fixed B - A -bimodule. Let $F: {}_A\mathcal{M} \rightarrow {}_B\mathcal{M}$ and $F': {}_B\mathcal{M} \rightarrow {}_A\mathcal{M}$ be the functors defined by

$$F(M) = L \otimes_A M, \quad F'(M') = \text{Hom}_B(L, M'),$$

where the assignment for morphisms is defined in the natural way. Then F is the left adjoint of F' and F' is the right adjoint of F . In particular, let $\rho: A \rightarrow B$ be a homomorphism and consider the case $L = B$. Then F is the functor $\rho^*: {}_A\mathcal{M} \rightarrow {}_B\mathcal{M}$ and F' is the functor $\rho_*: {}_B\mathcal{M} \rightarrow {}_A\mathcal{M}$, so that ρ^* is the left adjoint of ρ_* (and ρ_* is the right adjoint of ρ^*) (\rightarrow 277 Modules, K, L; for more examples of adjoint functors \rightarrow [11]).

L. Representation of Functors

We begin by discussing an example. Let T be a set, and consider the following problem: Is it possible to find a group X and a mapping $\xi: T \rightarrow X$ such that, for any group Y and for any mapping $\eta: T \rightarrow Y$, there exists a unique homomorphism $u: X \rightarrow Y$ with $u \circ \xi = \eta$? The answer is yes; it is enough to take the free group X generated by T and the canonical injection $\xi: T \rightarrow X$ (Fig. 6). On the other hand, let $F(Y)$ be the set of all mappings $T \rightarrow Y$, and for each group homomorphism $f: Y \rightarrow Y'$ define the mapping $F(f): F(Y) \rightarrow F(Y')$ by $F(f)\eta = f \circ \eta$ ($\eta \in F(Y)$). Then we get a covariant functor from the category \mathcal{C} of groups to the category of sets, $F: \mathcal{C} \rightarrow (\text{Sets})$. We can now reformulate the condition on $X \in \mathcal{C}$ and $\xi \in F(X)$ as follows:

For any $Y \in \mathcal{C}$ and for any $\eta \in F(Y)$, there exists a unique homomorphism $u: X \rightarrow Y$ such that $F(u)\xi = \eta$.



Fig. 6

Proceeding to the general case, let \mathcal{C} be an arbitrary category and let $F: \mathcal{C} \rightarrow (\text{Sets})$ be a functor. If there exist an object X of \mathcal{C} and an element ξ of $F(X)$ satisfying the condition just stated (with the modification $u: Y \rightarrow X$ in the contravariant case), then we say that the pair (X, ξ) **represents** the functor F , or less specifically, that X represents F , and we call ξ the **canonical element** of $F(X)$. We also say that F is **representable**. The condition stated above is a formulation of the so-called **universal map-**

ping property. If (X', ξ') also represents F , the unique morphism $u: X \rightarrow X'$ (or $X' \rightarrow X$) with $F(u)\xi = \xi'$ is necessarily an isomorphism.

When (X, ξ) represents F , the natural transformation $\varphi: h_X \rightarrow F$ ($h_X \rightarrow F$ in the contravariant case) which corresponds to ξ by the canonical bijection $\Phi_X: \text{Hom}(h_X, F) \xrightarrow{\sim} F(X)$ is an isomorphism. Conversely, if there is a functorial isomorphism $\varphi: h_X \rightarrow F$ (or $h^X \rightarrow F$) for some $X \in \mathcal{C}$, then the object X represents F , with the canonical element of $F(X)$ the element which corresponds to φ by the canonical bijection Φ_X , i.e., $\xi = \varphi(X)1_X$.

We have already seen the example of a free group; here we list a few more examples. (1) Let $\{X_i\}_{i \in I}$ be a family of objects in a category \mathcal{C} . For each $Y \in \mathcal{C}$ we put $F(Y) = \prod_{i \in I} \text{Hom}(Y, X_i)$, and for each morphism $f: Y \rightarrow Y'$ we define the mapping $F(f): F(Y') \rightarrow F(Y)$ by $F(f)(f_i) = (f_i \circ f)$. Then we get a contravariant functor $F: \mathcal{C} \rightarrow (\text{Sets})$. A pair (X, ξ) which represents F (where $\xi \in F(X) = \prod_{i \in I} \text{Hom}(X, X_i)$) is the product of $\{X_i\}$. Thus representability of F is equivalent to the existence of the product of $\{X_i\}$, and similarly for the coproduct.

(2) Let R be a ring, M a right R -module, and N a left R -module. For each Abelian group Y let $F(Y)$ denote the set of the R -balanced mappings $M \times N \rightarrow Y$ (\rightarrow 277 Modules J). Since a homomorphism $f: Y \rightarrow Y'$ induces a natural mapping $F(f): F(Y) \rightarrow F(Y')$ by composition, we obtain a covariant functor $F: (\text{Ab}) \rightarrow (\text{Sets})$. This functor is representable: the pair consisting of the tensor product $M \otimes_R N$ and the canonical mapping $M \times N \rightarrow M \otimes_R N$ represents it.

(3) Let R be a commutative ring and S a subset of R . For each commutative ring Y , let $F(Y)$ denote the set of homomorphisms $R \rightarrow Y$ that map the elements of S to invertible elements of Y . As in the preceding example, we obtain a covariant functor $F: (\text{Rings}) \rightarrow (\text{Sets})$. This functor is represented by the ring of quotients $S^{-1}R$ and the canonical homomorphism $R \rightarrow S^{-1}R$.

M. Groups in a Category

Let \mathcal{C} be a category with a final object e , and assume that a finite product always exists in \mathcal{C} . If an object $G \in \mathcal{C}$ and morphisms $\alpha: G \times G \rightarrow G$, $\beta: G \rightarrow G$, $\varepsilon: e \rightarrow G$ are given such that the diagrams of Fig. 7 are commutative, then $(G, \alpha, \beta, \varepsilon)$ is called a **group in \mathcal{C}** (**group object** in \mathcal{C} or \mathcal{C} -**group**).

If \mathcal{C} is the category of sets, then α defines a law of composition in the set G , and the image of e by ε is the identity element and $\beta(x)$ is the inverse of x , so that G is an ordinary group. If \mathcal{C} is the category of topological spaces (analy-

tic manifolds, algebraic varieties, ^{*}schemes) then G is a ^{*}topological group (^{*}Lie group, ^{*}algebraic group, ^{*}group scheme).

We can also define the \mathcal{C} -group by lifting the group concept in (Sets) to the category \mathcal{C} by means of the functor h^X . Namely, let G be an object of \mathcal{C} , and suppose that for each $Y \in \mathcal{C}$ the set $h^G(Y) = \text{Hom}(Y, G)$ is equipped with a group structure and that for each morphism $f: Y \rightarrow Y'$ the induced mapping $h^G(Y') \rightarrow h^G(Y)$ is a group homomorphism. In other words, suppose that h^G is a contravariant functor from \mathcal{C} to the category of groups. Then the object G with the additional structure on h^G is called a \mathcal{C} -group. This definition is equivalent to the one given above.

$$\begin{array}{ccc} G \times G \times G & \xrightarrow{\alpha \times 1_G} & G \times G \\ \downarrow 1_G \times \alpha & & \downarrow \alpha \\ G \times G & \xrightarrow{\alpha} & G \end{array}$$

$$\begin{array}{ccc} G \times G & \xrightarrow{1_G \times \epsilon} & G \times G \\ \downarrow p_1 & \swarrow \alpha & \downarrow \alpha \\ G & & G \end{array} \quad \begin{array}{ccc} G & \xrightarrow{(1_G, \beta)} & G \times G \\ \downarrow e & & \downarrow \alpha \\ G & \xrightarrow{\epsilon} & G \end{array}$$

Fig. 7

N. Additive Categories

A category \mathcal{C} is called an **additive category** if for each pair $X, Y \in \mathcal{C}$, the set of morphisms $\text{Hom}(X, Y)$ has the structure of an additive group such that (1) the composition of morphisms is distributive in both ways: $h \circ (f + g) = h \circ f + h \circ g$, $(f + g) \circ h = f \circ h + g \circ h$; (2) there exists an object $0'$ with $\text{Hom}(0', 0') = \{0\}$; (3) the product (or the coproduct) of any two objects exists. Then the object $0'$ in (2) is a final and initial object, and is called the **zero object**. Both the product and the coproduct of any two objects exist and can be identified. The dual category of an additive category is also an additive category. A functor F from an additive category to another is called an **additive functor** if $F(f + g) = F(f) + F(g)$ holds for morphisms. In an additive category \mathcal{C} , $\text{Hom}(X, Y)$ is an additive functor from \mathcal{C} to (Ab) in each variable.

For any ring R , the category of left (or right) R -modules is an additive category. The following definitions are generalizations of the corresponding concepts in the theory of modules. The **kernel** of a morphism $f: A \rightarrow B$ is a pair consisting of an object A' and a monomorphism $i: A' \rightarrow A$ with $f \circ i = 0$, such that any morphism $u: X \rightarrow A$ with $f \circ u = 0$ is divisible by i (that is, $u = i \circ v$ for some $v: X \rightarrow A'$). Dually, the **cokernel** of f is a pair consisting of an object B' and an epimorphism $j: B \rightarrow B'$ with $j \circ f = 0$ which divides any morphism $u: B \rightarrow X$ with $u \circ f = 0$. We write $A' = \text{Ker } f$, $B' =$

$\text{Coker } f$. The kernel of $j: B \rightarrow \text{Coker } f$ is called the **image** of f and is denoted by $\text{Im } f$; the cokernel of $i: \text{Ker } f \rightarrow A$ is called the **coimage** of f and is denoted by $\text{Coim } f$. If all these exist, it follows from the definitions that there is a unique morphism $\text{Coim } f \rightarrow \text{Im } f$ such that the composite of $A \rightarrow \text{Coim } f \rightarrow \text{Im } f \rightarrow B$ is equal to f .

An additive category \mathcal{C} is called an **Abelian category** if it satisfies the following conditions: (1) every morphism has a kernel and a cokernel, (2) for every morphism f , the morphism $\text{Coim } f \rightarrow \text{Im } f$ just mentioned is an isomorphism. The dual category of an Abelian category is also Abelian. The categories of Abelian groups, of R -modules, and of sheaves of \mathcal{O} -modules on a ^{*}ringed space (X, \mathcal{O}) are important examples of Abelian categories. Many propositions which are valid in (Ab) remain valid in any Abelian category. In particular, the notion of an ^{*}exact sequence is defined in an Abelian category in the same way as in (Ab), and the fiber product and fiber sum of a finite number of objects always exist in an Abelian category. A functor between Abelian categories which carries exact sequences into exact sequences is called an **exact functor**; (such a functor is automatically additive). If \mathcal{C} is a category of which $\text{Ob}(\mathcal{C})$ is a set and if \mathcal{C}' is an Abelian category, then $\text{Hom}(\mathcal{C}, \mathcal{C}')$ is an Abelian category. Given an Abelian category \mathcal{C} and a subcategory \mathcal{C}' which satisfies certain conditions, one can construct an Abelian category \mathcal{C}/\mathcal{C}' which is called the **quotient category** (Serre's theory of classes of Abelian groups; → [8]).

If \mathcal{C} is an Abelian category of which $\text{Ob}(\mathcal{C})$ is a set, there is an embedding of \mathcal{C} into the category ${}_R\mathcal{M}$ of modules over some ring R by a fully faithful flat exact covariant functor (**full embedding theorem**, B. Mitchell, *Amer. J. Math.*, 86 (1964)). This remarkable theorem enables us to extend results obtained for modules to the case of Abelian categories.

The notions of category and functor were introduced in [7] and were applied first in topology and then in homological algebra and algebraic geometry (→ 200 Homological Algebra).

References

- [1] S. Eilenberg and N. Steenrod, *Foundations of algebraic topology*, Princeton Univ. Press, 1952.
- [2] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton Univ. Press, 1956.
- [3] S. MacLane, *Homology*, Springer, 1963.

- [4] P. Freyd, Abelian categories, Harper & Row, 1964.
- [5] B. Mitchell, The theory of categories, Academic Press, 1965.
- [6] I. Bucur and A. Deleanu, Introduction to the theory of categories and functors, Wiley-Interscience, 1968.
- [7] S. Eilenberg and S. MacLane, General theory of natural equivalences, Trans. Amer. Math. Soc., 58 (1945), 231–294.
- [8] A. Grothendieck, Sur quelques points d'algèbre homologique, Tôhoku Math. J., (2) 9 (1957), 119–221.
- [9] P. Gabriel, Des catégories abéliennes, Thesis, Paris, 1962.
- [10] N. Yoneda, On universality I, II (in Japanese), Sûgaku, 13 (1961–1962), 109–112; 14 (1962–1963), 39–43.
- [11] S. MacLane, Categorical algebra, Bull. Amer. Math. Soc., 71 (1965), 40–106.

53 (XXI.17) Cauchy, Augustin Louis

The French mathematician Augustin Louis Cauchy (August 21, 1789–May 25, 1857) graduated from the Ecole Polytechnique in 1807 and from the Ecole des Ponts et Chaussées in 1810, to become a civil engineer. In 1816, his mathematical works were recognized, and he was appointed a member of the Académie des Sciences while a professor at the Ecole Polytechnique. After the July revolution in 1830, he refused to pledge loyalty to Louis-Philippe and fled to Turin; he later moved to Prague. He returned to France after the revolution of 1848 and became a professor at the University of Paris, where he remained until his death. He was a Catholic and a Royalist all his life.

His scientific contributions were numerous and covered many fields. In algebra, he did pioneering work in \dagger determinants and in the theory of \dagger groups. He also made notable achievements in theoretical physics, optics, and the theory of elasticity. His main field was analysis. He was interested in making analysis rigorous by giving calculus a solid foundation in such works as *Cours d'analyse de l'Ecole Polytechnique* (1821). In his paper “Memoire sur les intégrales définies prises entre les limites imaginaires” (1825), he proved the main theorem of the theory of functions of a complex variable. Another important work is his proof of the existence theorem for the solutions of \dagger differential equations in the cases of real and complex variables

References

- [1] A. Cauchy, Oeuvres complètes I.1–12; II.1–14, Gauthier-Villars, 1882–1958.
- [2] F. Klein, Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert I, Springer, 1926 (Chelsea, 1956).

54 (III.21) Cayley Algebras

Let Q be a \dagger quaternion algebra over a field K of characteristic zero. A **general Cayley algebra** \mathcal{C} is a 2-dimensional Q - \dagger module $Q + Qe$ with the multiplication $(q + re)(s + te) = (qs + \gamma \bar{t}r) + (tq + r\bar{s})e$, where $q, r, s, t \in Q$, γ is a given element in K , and \bar{t}, \bar{s} are the \dagger conjugate quaternions of t, s , respectively. The elements of \mathcal{C} are called **Cayley numbers**; \mathcal{C} is a nonassociative, \dagger alternative algebra of dimension 8 over K (\rightarrow 231 Jordan Algebras). The map $a = q + re \rightarrow \bar{a} = \bar{q} - re$ is an \dagger antiautomorphism of \mathcal{C} . Define two maps $\mathcal{C} \rightarrow K$ by $N(a) = a\bar{a} = \bar{a}a$ (**norm** of a) and $T(a) = a + \bar{a}$ (**trace** of a). Then every a in \mathcal{C} satisfies the equation $x^2 - T(a)x + N(a) = 0$. Furthermore, $N(ab) = N(a)N(b)$ for a, b in \mathcal{C} . The \dagger quadratic form $N(x) = T(x\bar{x})/2$ characterizes \mathcal{C} . In particular, any two (non-associative) general Cayley algebras over the same field K which are not \dagger alternative fields are isomorphic.

In order for \mathcal{C} to be an alternative field, either of the following two conditions is necessary and sufficient: (i) $N(a) = 0$ implies $a = 0$; (ii) Q is a noncommutative division algebra and γ cannot be expressed in the form $\sigma^2 - \lambda\xi^2 - \mu\eta^2 + \lambda\mu\zeta^2$ ($\sigma, \xi, \eta, \zeta \in K$). (For the meaning of λ, μ with respect to $Q \rightarrow$ 29 Associative Algebras D.) Every alternative field of finite dimension is a general Cayley algebra.

In particular, when Q is the \dagger quaternion field over the real number field with $\lambda = \mu = -1$, the general Cayley algebra over Q with $\gamma = -1$ is called the **Cayley algebra**. When K is an \dagger algebraic number field of finite degree, there are only a finite number of nonisomorphic general Cayley algebras over K .

The Lie algebra $\mathfrak{D}(\mathcal{C})$ of all \dagger derivations of a general Cayley algebra \mathcal{C} is a \dagger simple Lie algebra of type G_2 . If K is the real number field, the identity component of the group of all \dagger automorphisms of the Cayley algebra \mathcal{C} is a compact simply connected \dagger simple Lie group of type G_2 . The Cayley algebra \mathcal{C} is the unique alternative field over the real number field K . This last fact is important because of the following proposition: In the theory of \dagger non-

Desarguesian projective planes, the field which gives rise to the coordinates is an alternative field. Let \mathcal{P}_2 be the set of all 3×3 [†]Hermitian matrices A over the Cayley algebra \mathcal{C} such that $\text{tr } A = 1$, $A^2 = A$. Then we can define a structure of a projective plane on \mathcal{P}_2 , which with this structure is called the **Cayley projective plane**. Furthermore, let \mathfrak{J} be the set of all 3×3 [†]Hermitian matrices over \mathcal{C} , with a multiplication in \mathfrak{J} defined by $A \cdot B = (1/2)(BA + AB)$. The identity component G of the group of all automorphisms of \mathfrak{J} is a compact simply connected simple Lie group of type F_4 . This group G acts on \mathcal{P}_2 transitively, and $\mathcal{P}_2 = G/\text{Spin}(9)$ (\rightarrow 249 Lie Groups; Appendix A, Table 5.III).

References

- [1] L. E. Dickson, *Algebren und ihre Zahlen-theorie*, Orell Füssli, 1927.
- [2] M. Zorn, *Alternativkörper und quadratische Systeme*, Abh. Math. Sem. Univ. Hamburg, 9 (1933), 395–402.
- [3] N. Jacobson, Cayley numbers and normal simple Lie algebras of type G , *Duke Math. J.*, 5 (1939), 775–783.

55 (XX.6) Celestial Mechanics

A. General Remarks

The motions of planets, comets, the moon, and satellites in our solar system are the main topics in **celestial mechanics** (\rightarrow 392 Spherical Astronomy). However, studies in this subject can also include motions of fixed and binary stars in our galaxy, equilibrium figures of celestial bodies, and rotational motions of the earth and the moon.

Celestial mechanics is usually based on [†]Newtonian mechanics, with the effects of [†]general relativity sometimes taken into account to determine corrections to the orbits of celestial bodies. Therefore the main task of celestial mechanics is to solve differential equations of motion based on Newtonian mechanics. However, since the equations for the problem of $n > 2$ bodies cannot be solved rigorously (\rightarrow 420 Three-Body Problem), appropriate methods have been devised where by we may obtain approximate solutions of the equations with accuracy comparable to that of observations.

The **two-body problem**, which concerns the behavior of two celestial bodies regarded as

points exerting mutual interactions, can be reduced to a one-body problem with reference to a central force, since integrals of motion of the center of gravity for the system exist. The [†]Hamilton-Jacobi equation for the one-body problem is of [†]separable type and can be solved completely. The orbit for the two-body problem is a [†]conic with one of its [†]foci at the center of gravity. The majority of celestial bodies in the solar system actually perform **elliptic motions**. [†]Kepler's orbital elements for elliptic motion are functions of the integration constants in the solution of the Hamilton-Jacobi equation and are determined by the initial conditions.

B. Perturbations

In studying the [†] n -body problem, we first solve certain two-body problems and then apply the method of [†]perturbations, i.e., the **method of variation of constants**, in order to obtain solutions developed as [†]power series of small parameters. The parameters are ratios of the masses of planets to that of the sun for planetary motions and the ratio of the geocentric lunar distance to the solar distance for lunar motion.

Electronic computers have made it possible to compute planetary coordinates for long intervals of time by solving numerically differential equations of motion including all possible interactions. However, in discussing the stability of the solar system, analytic methods are more effective, particularly the method of obtaining **secular perturbations** by eliminating short-periodic terms by canonical transformations. This is one of the averaging methods of solving differential equations.

However, as the solutions obtained by means of perturbation methods are not always convergent, most important problems related to the stability of motion have not yet been solved rigorously. Secular perturbations for planetary motions can be derived by solving differential equations that are linearized by neglecting cubic powers of orbital eccentricities and inclinations to the ecliptic (\rightarrow 309 Orbit Determination), which are small quantities. The [†]eigenvalues of these linear differential equations correspond to mean angular velocities of the perihelion and the ascending node. The equations for the eigenvalues are called **secular equations**.

C. Artificial Satellites

To discuss motions of artificial satellites close to the earth, the latter cannot be regarded as a

point or as a sphere but must be assumed to be an oblate spheroid, i.e., an ℓ ellipsoid of revolution. The effects of oblateness on the motion of satellites can be derived as perturbations of the theoretical elliptic motions obtained as the solutions of this two-body problem under the assumption that the earth is spherical. Also, by utilizing a special potential very close to the geopotential, we can find a Hamilton-Jacobi equation of separable type which is solvable. This special potential appears in the problem of two fixed centers with equal masses situated on an imaginary axis. When the geopotential is assumed to be axially symmetric, the equations of motion for the satellites have two degrees of freedom; therefore there appear two fundamental frequencies related to the special potential. When these two frequencies are equal, the problem is called a **critical inclination problem** and is important from the mathematical point of view. Theories for satellites can be applied to the motions of fixed stars in the galaxy.

D. Equilibrium Figures

There is a large literature concerning **equilibrium figures** and the stabilities of celestial bodies assumed to consist of spinning fluids. The two-body problem with tidal interactions is particularly important; problems concerning the evolution of the earth-moon system are special cases of such a problem.

The theory of rotation of the earth as it is affected by \dagger precession, \dagger nutation, and latitude variations is also a part of celestial mechanics; for this theory, elastic theory and geophysics are applied.

For the n -body problem \rightarrow 420 Three-Body Problem.

References

- [1] F. F. Tisserand, *Traité de mécanique céleste* 1–4, Gauthier-Villars, 1889–1896.
- [2] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste* I, II, III, Gauthier-Villars, 1892–1899.
- [3] F. L. Moulton, *An introduction to celestial mechanics*, Macmillan, 1914.
- [4] A. Wintner, *The analytical foundation of celestial mechanics*, Princeton Univ. Press, 1941.
- [5] W. M. Smart, *Celestial mechanics*, Longmans-Green, 1953.
- [6] D. Brouwer and G. M. Clemence, *Methods of celestial mechanics*, Academic Press, 1961.
- [7] R. Deutsch, *Orbital dynamics of space vehicles*, Prentice-Hall, 1963.

- [8] Y. Hagihara, *Celestial mechanics. I, Dynamical principles and transformation theory*, MIT Press, 1970; *II, Perturbation theory*, pts. 1, 2, MIT, 1972; *III, Differential equations in celestial mechanics*, pts. 1, 2, Japan Society for the promotion of Science, 1974; *IV, Periodic and quadi-periodic solutions*, pts. 1, 2, Japan Society, 1975; *V, Topology of the three-body problem*, pts. 1, 2, Japan Society, 1976.
- [9] C. L. Siegel and J. K. Moser, *Lectures on celestial mechanics*, Springer, 1971.
- [10] E. L. Stiefel and G. Scheifele, *Linear and regular celestial mechanics*, Springer, 1971.

56 (IX.14) Characteristic Classes

A. General Remarks

The theory of characteristic classes arose from the problem of whether or not there exists a \dagger tangent r -frame field on a \dagger differentiable manifold (E. Stiefel [5]) (\rightarrow 105 Differentiable Manifolds). The importance of the characteristic class as a fundamental invariant of \dagger vector bundle structure is now fully recognized (\rightarrow 114 Differential Topology, 147 Fiber Bundles, 427 Topology of Lie Groups and Homogeneous Spaces).

B. Stiefel-Whitney Classes

Let $\xi = (E, B, \mathbf{R}^n)$ be an n -dimensional real \dagger vector bundle (called an \mathbf{R}^n -bundle) with \dagger paracompact Hausdorff \dagger base space B , fiber \mathbf{R}^n , and \dagger orthogonal group $O(n)$ as the \dagger structure group. Then the element $w_i(\xi)$ of the i -dimensional \dagger cohomology group $H^i(B; \mathbf{Z}_2)$ ($i = 1, 2, \dots, n$) of the base space B with coefficient in $\mathbf{Z}_2 = \mathbf{Z}/2\mathbf{Z}$, called the i -dimensional (or i th) **Stiefel-Whitney class**, and the element $w(\xi) = 1 + w_1(\xi) + \dots + w_n(\xi)$ of the \dagger cohomology ring $H^*(B; \mathbf{Z}_2)$, called the **total Stiefel-Whitney class**, are defined as follows. First, we deal with the case $n = 1$. We call the **infinite-dimensional real projective space** the \dagger inductive limit $\mathbf{P}^\infty(\mathbf{R}) = \varinjlim \mathbf{P}^n(\mathbf{R})$ of the finite-dimensional real \dagger projective space $\mathbf{P}^n(\mathbf{R})$. The nontrivial \dagger line bundle γ_1 over the infinite-dimensional real projective space $\mathbf{P}^\infty(\mathbf{R})$ is a universal \mathbf{R}^1 -bundle (\dagger universal bundle for the orthogonal group $O(1)$); therefore any \mathbf{R}^1 -bundle $\xi = (E, B, \mathbf{R}^1)$ is equivalent to an \dagger induced bundle from the universal bundle γ_1 by a \dagger classifying mapping $f_\xi: B \rightarrow \mathbf{P}^\infty(\mathbf{R})$: $\xi \equiv f_\xi^* \gamma_1$. We define the 1-dimensional universal Stiefel-Whitney class $w_1(\gamma_1)$ to be the generator of $H^1(\mathbf{P}^\infty(\mathbf{R}); \mathbf{Z}_2)$ and set $w_1(\xi) =$

$f_{\xi}^* w_1(\gamma_1)$. For general n , we consider the principal $O(n)$ -bundle $(P, B, O(n))$ \dagger associated with the given \mathbf{R}^n -bundle $\xi = (E, B, \mathbf{R}^n)$. Let Q_n be the subgroup of $O(n)$ consisting of all diagonal matrices. Then the orbit space $Y = P/Q_n$ is the base space of the \dagger principal Q_n -bundle $\eta = (P, P/Q_n, Q_n)$. Let $\rho: Y \rightarrow B = P/O(n)$ be the natural projection. Then the \mathbf{R}^n -bundle $\rho^* \xi$ over Y induced by ρ is associated with η and is equivalent to the \dagger Whitney sum of n line bundles (F. Hirzebruch [4]), $\rho^* \xi \equiv \xi_1 \oplus \dots \oplus \xi_n$. Moreover, the homomorphism $\rho^*: H^*(B; \mathbf{Z}_2) \rightarrow H^*(Y, \mathbf{Z}_2)$ is injective (A. Borel [1]). Therefore we can uniquely define the total Stiefel-Whitney class $w(\xi)$ of the \mathbf{R}^n -bundle by the relation $\rho^* w(\xi) = w(\xi_1) \dots w(\xi_n)$. The Stiefel-Whitney classes defined above are compatible with bundle mappings $f: w(f^* \xi) = f^* w(\xi)$ (\rightarrow 147 Fiber Bundles M). The Stiefel-Whitney class $w_i(\gamma_n) \in H^i(B_{O(n)}; \mathbf{Z}_2)$ ($1 \leq i \leq n$) of the universal \mathbf{R}^n -bundle γ_n over the \dagger classifying space $B_{O(n)}$ is called the i -dimensional **universal Stiefel-Whitney class**. For the Whitney sum of vector bundles, we have $w(\xi \oplus \eta) = w(\xi)w(\eta)$.

In order for an \mathbf{R}^n -bundle ξ to be \dagger orientable, namely, for ξ to have an $SO(n)$ -structure, it is necessary and sufficient that $w_1(\xi) = 0$. For an oriented \mathbf{R}^n -bundle ξ , the **Euler-Poincaré class** $X_n(\xi)$ is defined to be the \dagger primary obstruction $X_n(\xi) \in H^n(B; \mathbf{Z})$ for constructing a \dagger cross section of the \dagger associated $(n-1)$ -sphere bundle. In particular, the Euler-Poincaré class of the universal bundle for $SO(n)$ is called the **universal Euler-Poincaré class**. $X_n(\xi) \bmod 2$ is equal to $w_n(\xi)$. If n is odd, we have $2X_n(\xi) = 0$.

C. Chern Classes

We consider an n -dimensional complex vector bundle $\omega = (E, B, \mathbf{C}^n)$ (called \mathbf{C}^n -bundle in the following) with a paracompact Hausdorff base space B , fiber \mathbf{C}^n , and unitary group $U(n)$ as the structure group. The cohomology class $c_i(\omega) \in H^{2i}(B; \mathbf{Z})$ ($i = 1, 2, \dots, n$), called the $2i$ -dimensional (or i th) **Chern class**, and the **total Chern class** $c(\omega) = 1 + c_1(\omega) + \dots + c_n(\omega) \in H^*(B; \mathbf{Z})$ are defined as follows. In the case $n = 1$, let $\mathbf{C}^\infty = \varinjlim \mathbf{C}^n$ (the inductive limit of the \dagger complex Euclidean spaces \mathbf{C}^n), S^∞ be its unit sphere, and $\mathbf{P}^\infty(\mathbf{C})$ be the **infinite-dimensional complex projective space** consisting of all complex lines through the origin O of \mathbf{C}^∞ . Then the natural mapping $S^\infty \rightarrow \mathbf{P}^\infty(\mathbf{C})$ defines a universal principal $U(1)$ -bundle $(S^\infty, \mathbf{P}^\infty(\mathbf{C}), U(1))$. Let γ_1 be its associated universal \mathbf{C}^1 -bundle. Then we define the 1-dimensional universal Chern class $c_1(\gamma_1) \in H^2(\mathbf{P}^\infty(\mathbf{C}); \mathbf{Z})$ to be the cohomology class that takes the value -1 on the cycle $S^2 (\approx \mathbf{P}^1(\mathbf{C})) \subset$

$\mathbf{P}^\infty(\mathbf{C})$ with the natural orientation. Since a general \mathbf{C}^1 -bundle $\xi = (E, B, \mathbf{C}^1)$ is induced from γ_1 by a classifying mapping $f_\xi: B \rightarrow \mathbf{P}^\infty(\mathbf{C})$, we set $c_1(\xi) = f_\xi^* c_1(\gamma_1)$. When $n > 1$, let $(P, B, U(n))$ be the principal $U(n)$ -bundle associated with the given \mathbf{C}^n -bundle $\xi = (E, B, \mathbf{C}^n)$. Let T_n be the subgroup of $U(n)$ consisting of all diagonal matrices (which is a \dagger maximal torus of $U(n)$). Then the quotient space $Y = P/T_n$ is the base space of the principal T_n -bundle $\eta = (P, P/T_n, T_n)$. Let $\rho: Y \rightarrow B = P/U(n)$ be the natural projection. Then the \mathbf{C}^n -bundle $\rho^* \xi$ over Y is associated with η and is equivalent to the \dagger Whitney sum of n complex line bundles: $\rho^* \xi \equiv \xi_1 \oplus \dots \oplus \xi_n$. Moreover, $\rho^*: H^*(B; \mathbf{Z}) \rightarrow H^*(Y; \mathbf{Z})$ is a monomorphism (Borel [1], Hirzebruch [4]). Therefore we can uniquely define the total Chern class $c(\xi)$ of the \mathbf{C}^n -bundle ξ by the relation $\rho^* c(\xi) = c(\xi_1) \dots c(\xi_n)$. The Chern classes, as defined above, are compatible with bundle mappings f (i.e., $c(f^* \xi) = f^* c(\xi)$) (\rightarrow 147 Fiber Bundles N). The Chern class $c_i(\gamma_n) \in H^{2i}(B_{U(n)}; \mathbf{Z})$ ($1 \leq i \leq n$) of the universal \mathbf{C}^n -bundle γ_n over $B_{U(n)}$ is called the $2i$ -dimensional **universal Chern class**. Let ξ, η be complex vector bundles over B . Then we have $c(\xi \oplus \eta) = c(\xi)c(\eta)$. By the natural inclusion $U(n) \subset SO(2n)$, we can identify a \mathbf{C}^n -bundle ω with an oriented \mathbf{R}^{2n} -bundle $\omega_{\mathbf{R}}$. Then we have $c_i(\omega) \bmod 2 = w_{2i}(\omega_{\mathbf{R}})$, $w_{2i+1}(\omega_{\mathbf{R}}) = 0$ ($i = 0, 1, \dots, n$), $c_n(\omega) = X_{2n}(\omega_{\mathbf{R}})$.

Examples. Let $(S^{2n+1}, \mathbf{P}^n(\mathbf{C}), U(1))$ be the \dagger Hopf bundle, and let γ_1^n be the associated complex line bundle. Then the classifying mapping of γ_1^n is the natural inclusion $\mathbf{P}^n(\mathbf{C}) \rightarrow BU(1) = \mathbf{P}^\infty(\mathbf{C})$, and $c(\gamma_1^n) = 1 - g_n$, where $g_n \in H^2(\mathbf{P}^n(\mathbf{C}); \mathbf{Z})$ is the cohomology class dual to the homology class represented by the hyperplane $\mathbf{P}^{n-1}(\mathbf{C})$. On the other hand, for the complex line bundle $\xi_1^n = \{\mathbf{P}^{n-1}(\mathbf{C})\}$ determined by the \dagger divisor (\rightarrow 72 Complex Manifolds) $\mathbf{P}^{n-1}(\mathbf{C}) \subset \mathbf{P}^n(\mathbf{C})$, we have $c(\xi_1^n) = 1 + g_n$, and ξ_1^n, γ_1^n are dual to each other. ξ_1^n is called the **canonical line bundle** over the complex projective space. Moreover, the Whitney sum $\tau \oplus \varepsilon_1$ of the tangent bundle $\tau(\mathbf{P}^n(\mathbf{C}))$ of the n -dimensional complex projective space $\mathbf{P}^n(\mathbf{C})$ and the trivial \mathbf{C}^1 -bundle ε_1 is equivalent to the Whitney sum of $(n+1)$ copies of ξ_1^n . Therefore we have $c(\tau(\mathbf{P}^n(\mathbf{C}))) = c(\xi_1^n \oplus \xi_1^n \oplus \dots \oplus \xi_1^n) = (1 + g_n)^{n+1}$.

D. Pontryagin Classes

Utilizing the inclusion mapping $O(n) \subset U(n)$ of structure groups, we can associate a \mathbf{C}^n -bundle $\xi_{\mathbf{C}} = \xi \oplus \sqrt{-1} \xi$ with an \mathbf{R}^n -bundle ξ . We define the $4i$ -dimensional (or i th) **Pontryagin class** of the \mathbf{R}^n -bundle ξ by $p_i(\xi) = (-1)^i c_{2i}(\xi_{\mathbf{C}})$ ($\in H^{4i}(B; \mathbf{Z})$) ($i = 1, 2, \dots, [n/2]$) (Hirzebruch

[4]). In particular, the Pontryagin classes of the universal bundle for $O(n)$ are called the **universal Pontryagin classes**. We also define the **total Pontryagin classes** $p(\xi) = 1 + p_1(\xi) + \dots + p_{[n/2]}(\xi)$. We have $2c_{2i+1}(\xi_C) = 0$. For \mathbf{R}^n -bundles ξ and η , $p(\xi \oplus \eta) = p(\xi)p(\eta)$ modulo 2-torsion elements. We have $p_i(\xi) \pmod{2} = (w_{2i}(\xi))^2$ and $p_n(\xi) = (X_{2n}(\xi))^2$ for the oriented \mathbf{R}^{2n} -bundle ξ . Moreover, for a complex vector bundle ω , we have (Wu [8])

$$(-1)^k p_k(\omega_{\mathbf{R}}) = \sum_{i=0}^{2k} (-1)^i c_i(\omega) c_{2k-i}(\omega),$$

where we put $c_0(\omega) = 1$.

All such classes as defined in Sections B–D are called **characteristic classes**.

E. Other Definitions of Characteristic Classes

Axiomatic Definition. (1) For a \mathbf{C}^n -bundle ξ over a paracompact Hausdorff base space B , Chern classes $c_i(\xi) \in H^{2i}(B; \mathbf{Z})$ ($i \geq 0$) are defined, and we have $c_0(\xi) = 1$, $c_i(\xi) = 0$ ($i > n$). (2) For the total Chern class $c(\xi) = \sum_{i=0}^{\infty} c_i(\xi)$, we have $c(f^* \xi) = f^* c(\xi)$ for each bundle mapping f . (3) For the Whitney sum, we have $c(\xi \oplus \eta) = c(\xi) \cdot c(\eta)$. (4) Normalization condition: For the canonical line bundle ξ_1^n , we have $c(\xi_1^n) = 1 + g_n$ (\rightarrow Section C). We can verify the existence and uniqueness of $c_i(\xi)$ satisfying these four conditions, so Chern classes can be defined axiomatically by these conditions (Hirzebruch [4]). We can similarly define Stiefel-Whitney classes axiomatically.

Definition in Terms of Obstruction Classes.

When the base space is a \dagger CW complex, we can define the Chern class of a \mathbf{C}^n -bundle ξ over B as follows: Let $V_{n,n-q+1}(\mathbf{C}) = U(n)/I_{n-q+1} \times U(q-1)$ be the \dagger complex Stiefel manifold of all orthonormal $(n-q+1)$ -frames in \mathbf{C}^n with Hermitian metric. Then $V_{n,n-q+1}(\mathbf{C})$ is $\dagger(2q-2)$ -connected, and its $(2q-1)$ -dimensional \dagger homotopy group $\pi_{2q-1}(V_{n,n-q+1}(\mathbf{C})) = \mathbf{Z}$. Let ξ' be the \dagger associated bundle of ξ with fiber $V_{n,n-q+1}(\mathbf{C})$. Then the \dagger primary obstruction ($\in H^{2q}(B; \mathbf{Z})$) to constructing a \dagger cross section of ξ' coincides with the Chern class $c_q(\xi)$. Analogously, we can interpret $w_q(\xi)$ for an \mathbf{R}^n -bundle ξ as an obstruction class (\rightarrow 147 Fiber Bundles).

Definition in Terms of Schubert Varieties.

We denote by \mathbf{C}^k the subspace defined by $z_{k+1} = z_{k+2} = \dots = z_{n+N} = 0$ of the space $\mathbf{C}^{n+N} = \{(z_1, \dots, z_{n+N}) \mid z_i \in \mathbf{C}, i = 1, \dots, n+N\}$, and fix the sequence of subspaces $\mathbf{C}^1 \subset \mathbf{C}^2 \subset \dots \subset \mathbf{C}^{n+N}$. The set of all complex n -planes X through the origin O in \mathbf{C}^{n+N} forms the \dagger complex Grassmann manifold $M_{n+N,n}(\mathbf{C})$. We denote

by $E(\gamma_n^N)$ the set of all pairs (X, v) , where $X \in M_{n+N,n}(\mathbf{C})$ and v is a vector in X . Then we can define a $2N$ -universal complex n -dimensional vector bundle γ_n^N with base space $M_{n+N,n}(\mathbf{C})$, with total space $E(\gamma_n^N)$ and projection $(X, v) \rightarrow X$. Let $\omega = (\omega(1), \dots, \omega(n))$ be a sequence of integers satisfying the condition $0 \leq \omega(1) \leq \dots \leq \omega(n) \leq N$. Then the set e_ω of all n -planes $X \subset \mathbf{C}^{n+N}$ through the origin O satisfying $\dim(X \cap \mathbf{C}^{i+\omega(i)}) = i$, $\dim(X \cap \mathbf{C}^{i+\omega(i)-1}) = i-1$, $i = 1, 2, \dots, n$, forms a real $(2 \sum_{i=1}^n \omega(i))$ -dimensional open \dagger cell. The set of all these open cells e_ω gives a \dagger cellular subdivision of $M_{n+N,n}(\mathbf{C})$ as a CW complex. The closure \bar{e}_ω of e_ω is a cellular subcomplex of $M_{n+N,n}(\mathbf{C})$, called a **Schubert variety**. This is a \dagger pseudo-manifold with canonical orientation and represents a $(2 \sum_{i=1}^n \omega(i))$ -dimensional integral cycle, called a **Schubert cycle**. We denote \bar{e}_ω by $(\omega(1), \dots, \omega(n))$. All these homology classes form the basis of the homology group $H_*(M_{n+N,n}(\mathbf{C}); \mathbf{Z})$. The cocycle dual to the cycle $(\underbrace{0, \dots, 0}_{n-q}, \underbrace{1, \dots, 1}_q)$ represents the Chern class $c_q(\gamma_n^N) \in H^{2q}(M_{n+N,n}(\mathbf{C}); \mathbf{Z})$. For the real Grassmann manifold $M_{n+N,n}(\mathbf{R})$, we can analogously define the universal Stiefel-Whitney classes.

Thom's Definition. Let ξ be an \mathbf{R}^n -bundle over B , B_ξ be its \dagger Thom space, and $U \in H^n(B_\xi; \mathbf{Z}_2)$ be the \dagger fundamental class of B_ξ . Let $j: B \rightarrow B_\xi$ be the inclusion induced from the zero cross section and $\varphi: H^k(B; \mathbf{Z}_2) \cong H^{k+n}(B_\xi; \mathbf{Z}_2)$ be the \dagger Thom-Gysin isomorphism. Then we have $j^* U = w_n(\xi)$, $\varphi^{-1}(Sq^i U) = w_i(\xi)$ ($0 \leq i \leq n$), where Sq^i is the \dagger Steenrod square (R. Thom [6]).

Definition in Terms of Differential Forms.

Let B be a \dagger differentiable manifold and $\xi = (P_\xi, B, U(n))$ be a differentiable principal $U(n)$ -bundle over B . Let $\Omega = (\Omega_{ij})$ be the \dagger curvature form corresponding to the \dagger connection form $\omega = (\omega_{ij})$, $i, j = 1, \dots, n$, on P_ξ . Then Ω_{ij} is a complex-valued 2-form, and $\bar{\Omega}_{ij} = -\Omega_{ji}$. For the matrix Ω , we consider the differential form

$$\psi = \sum_q \psi_q = \det[I + (2\pi\sqrt{-1})^{-1} \Omega],$$

where I is the unit matrix, the multiplication in the determinant is the \dagger exterior product, and ψ_q is the part of degree $2q$ in ψ . Then ψ is defined as a real form independent of the connection ω . We have $d\psi_q = 0$, and the cohomology class of $(-1)^q \psi_q$ in $H^{2q}(B; \mathbf{R})$ is the Chern class $c_q(\xi)$ with real coefficients (Borel and Hirzebruch [2], S. S. Chern [3]).

Definition in Terms of Symmetric Polynomials.

(\rightarrow 427 Topology of Lie Groups and Homogeneous Spaces.)

F. Characteristic Classes of Manifolds

For a differentiable (complex or almost complex) manifold M , the characteristic classes of its tangent bundle are called **characteristic classes of the manifold** M . We shall denote **Stiefel-Whitney classes**, **Pontryagin classes**, **Euler-Poincaré classes**, and **Chern classes** of M by $w_i(M)$, $p_i(M)$, $X_n(M)$, and $c_i(M)$, respectively. These are invariants of differentiable structures, orientations, or (almost) complex structures of a manifold M if M is a differentiable, oriented differential, or (almost) complex manifold. By the **Stiefel-Whitney numbers** of an n -dimensional manifold M , we mean the values of n -dimensional monomials of Stiefel-Whitney classes of M on the fundamental homology class $((w_1(M)^{r_1} w_2(M)^{r_2} \dots w_n(M)^{r_n})[M] \in \mathbb{Z}_2$, where $r_1 + 2r_2 + \dots + nr_n = n$, $r_i \geq 0$). We can define integer-valued **Pontryagin numbers** and **Chern numbers** similarly. These numbers are called generally **characteristic numbers** of the given manifold. In particular, $X_n(M)[M] = \chi(M)$ is the \dagger Euler-Poincaré characteristic.

In the case of topological manifolds, we can define characteristic classes in the following sense. Let M be a closed n -dimensional topological manifold and X^n the generator of $H^n(M; \mathbb{Z}_2)$. By defining $X^i(Y^{n-i}) = (X^i Y^{n-i})[M] \in \mathbb{Z}_2$ for $X^i \in H^i(M; \mathbb{Z}_2)$, $Y^{n-i} \in H^{n-i}(M; \mathbb{Z}_2)$, we have an isomorphism $H^i(M; \mathbb{Z}_2) \cong \text{Hom}(H^{n-i}(M, \mathbb{Z}_2); \mathbb{Z}_2)$. The element $u_i \in H^i(M; \mathbb{Z}_2)$, corresponding to the homomorphism $Y^{n-i} \rightarrow Sq^i Y^{n-i}[M]$ under this isomorphism is called the **Wu class** of M , where Sq^i is the \dagger Steenrod square. Moreover, we call $w_j = \sum_{i=0}^j Sq^{j-i} u_i \in H^j(M; \mathbb{Z}_2)$ the **Stiefel-Whitney class** of the topological manifold M . Then for any \dagger differentiable structure \mathcal{D} , we have $w_j(M, \mathcal{D}) = w_j$. Therefore Stiefel-Whitney classes of differentiable manifolds are topological invariants (more precisely homotopy type invariants) (Thom [6], W. T. Wu [8]). J. W. Milnor [9] proved that Pontryagin classes of differentiable manifolds are not topological invariants. The image of $p_i(M)$ by the homomorphism $H^{4i}(M; \mathbb{Z}) \rightarrow H^{4i}(M; \mathbb{Q})$ induced by the inclusion $\mathbb{Z} \subset \mathbb{Q}$ (the rational number field) is called the **rational Pontryagin class**. In 1966, S. P. Novikov [10] proved the topological invariance of the rational Pontryagin class (\rightarrow Section H).

G. Index Theorem for Differentiable Manifolds

Let M be an oriented closed manifold of dimension $4k$. Putting $f(x, y) = x \cdot y[M]$ for elements x, y of the $2k$ -dimensional real coho-

mology group $H^{2k}(M; \mathbb{R})$, we obtain a bilinear form on $H^{2k}(M; \mathbb{R})$. The \dagger signature of the quadratic form $f(x, x)$ (namely, (number of positive terms) – (number of negative terms) in its canonical form) is a topological invariant (homotopy type invariant) of the manifold M . We call it the **index** or the **signature** of the manifold M and denote it by $\tau(M)$. If the dimension of M is not divisible by 4, we define $\tau(M) = 0$. For the product of manifolds we have $\tau(M \times N) = \tau(M) \cdot \tau(N)$. Also $\tau(M)$ is an invariant of the \dagger cobordism class of M (Thom [6]).

The index τ of a differentiable manifold gives a homomorphism of the \dagger cobordism ring Ω into the ring \mathbb{Z} of integers. Hirzebruch investigated the multiplicative property of τ and gave its expression by means of Pontryagin numbers. Let P_i be the i th \dagger elementary symmetric function of indeterminates β_1, \dots, β_n . Then a homogeneous part of the formal power series

$$\prod_{i=1}^n \frac{\sqrt{\beta_i}}{\tanh \sqrt{\beta_i}}$$

of β_1, \dots, β_n is a symmetric polynomial of β_1, \dots, β_n , and therefore a polynomial of P_i with rational coefficients. For $k \leq n$, we denote the homogeneous part of degree k by $L_k(P_1, \dots, P_k)$. Specifically, if P_i are the Pontryagin classes $p_i(M^{4k})$ of a $4k$ -dimensional closed differentiable manifold M^{4k} , then $L_k(P_1, \dots, P_k)$ is a $4k$ -dimensional cohomology class of M^{4k} . Then we have the formula

$$\tau(M^{4k}) = L_k(P_1, \dots, P_k)[M^{4k}],$$

called the **index theorem** of differentiable manifolds (or **Hirzebruch index theorem**). For example, $L_1 = (1/3)P_1$, $L_2 = (1/45)(7P_2 - P_1^2)$, and $L_3 = (1/945)(62P_3 - 13P_2P_1 + 2P_1^3)$, \dots (\rightarrow 114 Differential Topology). Later this index theorem was generalized to the \dagger Atiyah-Singer index theorem (\rightarrow 237 K-Theory).

H. Combinatorial Pontryagin Classes

Let K be an oriented n -dimensional compact \dagger homology manifold, and let Σ^r be the boundary of an oriented $(r+1)$ -simplex, namely, the combinatorial r -sphere. Let $f: K \rightarrow \Sigma^{n-4i}$ be a \dagger piecewise linear mapping. Then for almost all points y of Σ^{n-4i} , $f^{-1}(y)$ is an oriented $4i$ -dimensional compact homology manifold, and its index $\tau(f^{-1}(y))$ is independent of y . We denote this by $\tau(f)$. The $\tau(f)$ is an invariant of the homotopy class of f . Let σ be the fundamental class of $H^{n-4i}(\Sigma^{n-4i}; \mathbb{Z})$. Then for $n \geq 8i + 2$, there exists a unique cohomology class $l_i = l_i(K) \in H^{4i}(K; \mathbb{Q})$ such that for any piecewise linear mapping $f: K \rightarrow \Sigma^{n-4i}$, we have

$(l_i \cdot f^* \sigma)[K] = \tau(f)$. We can remove the restriction $n \geq 8i + 2$ if we take $K \times \Sigma^m$ for K and define $l_i(K)$ to be $l_i(K \times \Sigma^m)$ for sufficiently large m . If K is a $\dagger C^1$ -triangulation of a differentiable manifold M , $l_i(K)$ coincides with the class $L_i(P_1, \dots, P_j)$ defined by Hirzebruch (R. Thom [7]; V. Rokhlin and A. Shvarts), where $P_j = p_j(M)$ is the Pontryagin class of M . Since the variable P_i can be expressed as a polynomial with rational coefficients of $L_j(P_1, \dots, P_j)$, $j \leq i$, we define the **combinatorial Pontryagin class** $p_i(K)$ of a homology manifold K as the polynomial of $l_j(K)$ with rational coefficients. Therefore, if K is a C^1 -triangulation of a differentiable manifold M , we have $p_i(K) = p_i(M)$ (modulo torsion elements). The class $l_i(K)$ and consequently $p_i(K)$ are important combinatorial invariants of K . These classes are topological invariants [10].

References

- [1] A. Borel, Sur la cohomologie des espaces fibrés principaux et des espaces homogènes de groupes de Lie compacts, *Ann. Math.*, (2) 57 (1953), 115–207.
- [2] A. Borel and F. Hirzebruch, Characteristic classes and homogeneous spaces I, II, III, *Amer. J. Math.*, 80 (1958), 459–538; 81 (1959), 315–382; 82 (1960), 491–504.
- [3] S. S. Chern, Characteristic classes of Hermitian manifolds, *Ann. Math.*, (2) 47 (1946), 58–121.
- [4] F. Hirzebruch, *Topological methods in algebraic geometry*, Springer, third edition, 1966.
- [5] E. Stiefel, Richtungsfelder und Fernparallelismus in n -dimensionalen Mannigfaltigkeiten, *Comment. Math. Helv.*, 8 (1936), 3–51.
- [6] R. Thom, Espaces fibrés en sphères et carrés de Steenrod, *Ann. Sci. Ecole Norm. Sup.*, (3) 69 (1952), 109–182.
- [7] R. Thom, Les classes caractéristiques de Pontrjagin des variétés triangulées, *Symposium Internacional de Topologia Algebraica*, Mexico, 1958.
- [8] W. T. Wu, Sur les classes caractéristiques des structures fibrés sphériques, *Actualités Sci. Ind.*, Hermann, 1952.
- [9] J. W. Milnor, Microbundles I, *Topology*, 3 (1964), Suppl. I., 53–80.
- [10] S. P. Novikov, Manifolds with free Abelian fundamental groups and their applications (Pontrjagin classes, smoothness, multi-dimensional knots), *Amer. Math. Soc. Transl.*, 71 (1968), 1–42. (Original in Russian, 1966.)
- [11] N. Steenrod, *The topology of fibre bundles*, Princeton Univ. Press, 1951.
- [12] D. Husemoller, *Fibre bundles*, McGraw-Hill, 1966.
- [13] J. Milnor and J. Stasheff, *Characteristic classes*, Princeton Univ. Press, 1974.

57 (XXI.6) Chinese Mathematics

A. Mathematics in the Chao, Han, and Tang Dynasties (3rd Century B.C.–10th Century A.D.)

In ancient China, the art of divination, called *yi*, was used in government administration. This was a kind of calculation that used pieces called *tse*. The book embodying it, called the *I-Ching*, is still popularly used. It shows that “numbers” or mathematics was seriously utilized in China at that time. The multiplication table for numbers up to nine (called the Pythagorean table in the West) was known in China from the legendary period. However, mathematics in the Greek sense, that is, mathematics as a logically systematized science, was unknown in ancient China.

Suanching-Shihshu, or the Ten Books on Arithmetic—namely, *Choupi-Suanching*, *Chiuchang-Suanshu*, *Haitao-Suanching* (edited by Liu Hui), *Suntzu-Suanching*, *Wutsao-Suanching*, *Hsiahouyang-Suanching*, *Changchiu-Suanching*, *Wuching-Suanshu*, *Chiku-Suanching* (edited by Wang Hsiao-Tong), and *Shushu-Chiyi* (edited by Hsu Yue)—came into being between the 2nd century B.C. and the 6th century A.D., from the Chao to the Han eras, with the exception of the *Chiku-Suanching* compiled in the Tang era. These are the only mathematical texts from this early period whose authors and times of publication are known. They were used in the civil service examination for selecting administrators up to the beginning of the Sung era (960 A.D.). The most important among them is *Chinchang-Suanshu*, or the Book of Arithmetic, which contains nine chapters. It treats positive and negative fractions with laws of operations on signed numbers, equations, and the elementary mathematical knowledge of daily life. The *Chiku-Suanching* contains a number of problems reducible to equations of the 3rd and 4th degrees.

There were also two works called *San Tung Shu* (edited by Tong Chuan) and *Chui Shu* (edited by Tsu Chung-Chih), but no copies of them are extant. Later works, one from the Sui era (published in 636) and another from the Tang era, tell us that the latter contained the result $3.1415927 > \pi > 3.1415926$ and the approximate values $355/113$ and $22/7$ for π .

In the 1st century A.D. Buddhism was introduced from India, and paper was invented.

However, despite the communication with India, neither the Indian numeration system, written calculation, nor the abacus was at that time widely used in China. The extraction of square or cube roots was done with calculating rods.

B. Mathematics in the Sung and Yuan Dynasties (10th–14th Centuries)

In the Sung and Yuan periods contact was made with the Arab world. In the 13th century, a mechanical algebra utilizing calculating rods made remarkable progress; this can be attributed to Arab influence. Toward the end of the Sung era appeared the *Shushu-Chiuchang* by Ch'in Chiu-Shao and the *Yiku-Yentan* by Li Chih. The former gives a method like Horner's for approximate solution of equations, and the latter gives the principle of *tiennyuan-shu*, i.e. the mechanical algebra of this period. The principle of *tiennyuan-shu* was further expounded in the *Suanhsueh Chimeng* (1295) and the *Suyuan Yuchien* (1303) by Shih Shih-Chieh, the *Yanghui Suanfa* by Yang Hui, and other works. These were introduced into Japan and they influenced the *wasan* (Japanese mathematics) of early times. Until recently, no further original mathematical ideas appeared in China.

C. Mathematics after the Ming Era (15th Century)

In this epoch, European renaissance civilization began to influence the Orient. In 1607, Matteo Ricci (1552–1610) translated Books I–IV of Euclid's *Elements* into Chinese with the aid of Hsu Kuang-Chi. In 1592 *Suanfa Tangtsung* by Ch'êng Ta-Wei appeared, which dealt with the use of the abacus. This book had great influence upon *wasan*.

No development was seen in the indigenous mathematics of the Ching era, that is, after the 17th century, but science and technology were imported by Christian missionaries. This brought about calendar reform from the lunar to the solar method. On the other hand, new editions of classical works such as the Ten Books on Arithmetic began to appear in this period. Emperor Kang Hsi-Ti (1655–1722), who was in correspondence with Leibniz, asked Ferdinand Verbiest (renamed Nan Huai Jen in Chinese) to compile *Shuli-Chingwen* (Advanced mathematics), whose 53 chapters were completed in 1723. This book dealt with European-style algebra and trigonometry. In the latter half of the 19th century, Alexander Wylie translated a number of Western mathe-

matical books into Chinese, including Books VII–XIII of Euclid's *Elements* and some works on calculus. Many current Chinese and Japanese mathematical terms originated with this translation.

References

- [1] Y. Li, History of mathematics in China (Japanese translation), Seikatusya, 1940.
- [2] Y. Mikami, The development of mathematics in China and Japan, Teubner, 1913 (Chelsea, 1961).

58 (X.9) C^∞ -Functions and Quasi-Analytic Functions

A. General Remarks

An example of a ${}^*C^\infty$ -function is a * real analytic function, which is defined to be a function that can be expressed as a power series that converges in a neighborhood of each point of the domain where the function is defined. Many examples, however, show that real analytic functions form a rather small subset of the C^∞ -functions. Sometimes C^∞ -functions (not real analytic functions) play essential roles in the development of theories of analysis (\rightarrow 105 Differentiable Manifolds S). On the other hand, there is a subfamily of C^∞ -functions having some remarkable properties in common with the family of real analytic functions. This family is called the **family of quasi-analytic functions**. It has been an important object of study since the beginning of the 20th century. On the other hand, the Gevrey class functions, which are no longer quasi-analytic in general, also constitute an important subset of C^∞ -functions. The first part of this article deals with C^∞ -functions, the second part with quasi-analytic functions, and the third part with Gevrey class functions.

B. C^∞ -Functions

Let Ω be an open set of the n -dimensional real Euclidean space \mathbf{R}^n . A real-valued function $f(x_1, \dots, x_n)$ defined on Ω is called a **function of class C^∞** on Ω (or C^∞ -function on Ω) if $f(x_1, \dots, x_n)$ is continuously differentiable up to any order. The totality of C^∞ -functions defined on Ω is denoted by $C^\infty(\Omega)$. It is an * associative algebra over the real number field \mathbf{R} . A continuous function f defined on some closed set F

of \mathbf{R}^n is called a C^∞ -function on F if there exist an open neighborhood U of F and a function $g \in C^\infty(U)$ such that $f = g|_F$. This definition is equivalent to the following (H. Whitney [6]): For any multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$, we can find a continuous function $f^\alpha(x_1, \dots, x_n)$ on F such that (i) $f^0(x_1, \dots, x_n) = f(x_1, \dots, x_n)$ and (ii) for any positive integer r and for every multi-index α with $|\alpha| \leq r$,

$$\lim_{y \rightarrow x} \left[\frac{1}{\|y - x\|^{r-|\alpha|}} \left| f^\alpha(x) - \sum_{|\alpha+\beta| \leq r} f^{\alpha+\beta}(x) \frac{(x-y)^\beta}{\beta!} \right| \right] = 0,$$

where $\| \cdot \|$ denotes the Euclidean norm of \mathbf{R}^n .

C. Local Theory of C^∞ -Functions

We shall now introduce an equivalence relation \sim in $C^\infty(\mathbf{R}^n)$, defined as follows: $f \sim g \Leftrightarrow f|U = g|U$ for some open neighborhood U of the origin. Let \mathcal{E}_n denote the quotient set $C^\infty(\mathbf{R}^n)/\sim$, which naturally inherits the structure of an associative algebra from $C^\infty(\mathbf{R}^n)$. An element of \mathcal{E}_n is called a **germ of a C^∞ -function at the origin**. We denote the germ of $f \in C^\infty(\mathbf{R}^n)$ by \tilde{f} . To $\tilde{f} \in \mathcal{E}_n$, we assign the **formal Taylor expansion** $\sum (D^\alpha f(0)/\alpha!) x^\alpha$ around 0, where $D^\alpha f$ means $(\partial^{r_1} + \dots + \partial^{r_n})f / (\partial^{r_1} x_1 \dots \partial^{r_n} x_n)$ for $\alpha = (r_1, \dots, r_n)$. This assignment induces a homomorphism τ from \mathcal{E}_n to the ring of formal power series $\mathbf{R}[[x_1, \dots, x_n]]$ of n variables. The homomorphism τ is surjective but not injective. Put $\Lambda_n = \tau^{-1}(0) \subset \mathcal{E}_n$. A function f whose germ \tilde{f} belongs to Λ_n is called a **flat function**. The function $\varphi(x)$ defined by $\varphi(x) = \exp(-1/x^2)$ when $x \neq 0$ and $\varphi(0) = 0$ is an example of a flat function on \mathbf{R}^1 . A close study of the relationship between \mathcal{E}_n and $\mathbf{R}[[x_1, \dots, x_n]]$ leads to the **preparation theorem for C^∞ -functions**, which can be stated as follows: Let $\tilde{F}(x_1, \dots, x_n) \in \mathcal{E}_n$ satisfy $\tilde{F}(0, \dots, 0, x_n) = x_n^p \tilde{g}(x_n)$ ($\tilde{g} \in \mathcal{E}_1, g(0) \neq 0$). Then any $\tilde{f} \in \mathcal{E}_n$ can be expressed as $\tilde{f} = \tilde{F} \tilde{Q} + \tilde{R}$, where $\tilde{Q} \in \mathcal{E}_n$ and $\tilde{R} = \sum_{i=0}^{p-1} r_i(x_1, \dots, x_{n-1}) x_n^i$ with $r_i \in \mathcal{E}_{n-1}$ (B. Malgrange [3]).

Let $f(x_1, \dots, x_n)$ be a symmetric function in (x_1, \dots, x_n) of class C^∞ . Then there exists a germ $\tilde{g} \in \mathcal{E}_1$ such that $\tilde{f}(x_1, \dots, x_n) = \tilde{g}(\sigma_1, \dots, \sigma_n)$, where $\sigma_1, \dots, \sigma_n$ denote elementary symmetric functions with respect to x_1, \dots, x_n (G. Glaeser, Malgrange). Let $\tilde{f} \in \mathcal{E}_1$ satisfy $f(x) = f(-x)$. Then there exists a germ $\tilde{g} \in \mathcal{E}_1$ such that $\tilde{f}(x) = \tilde{g}(x^2)$ (H. Whitney).

D. Global Results

Case of n Variables. $C^\infty(\Omega)$ becomes a * Fréchet space when it is endowed with the topology of

uniform convergence on compact sets for all partial derivatives. Let J and J_1 be two closed * ideals of $C^\infty(\Omega)$. Then we have $J = J_1$ if and only if $\tau_x(J) = \tau_x(J_1)$ for each $x \in \Omega$, where τ_x is the mapping from $C^\infty(\Omega)$ to the ring of the formal power series that assigns the formal Taylor series of f around x (Whitney).

Case of One Variable. In the case of one variable, further information can be obtained from various points of view. In the following, f denotes a C^∞ -function defined on the unit interval $I = [0, 1]$. If f satisfies $f(1) = 1, f'(0) = \dots = f^{(r-1)}(0) = 0$, then we have

$$\frac{1}{m_1} + \frac{1}{\sqrt{m_2}} + \dots + \frac{1}{\sqrt[r]{m_r}} < k,$$

where $m_i = \sup\{|f^{(i)}(x)| \mid x \in I\}$ and where k is some constant independent of the choice of f and r (E. Borel). Similar kinds of inequalities were obtained by A. N. Kolmogorov, A. Gorny, and H. Cartan. Let A be an arbitrary countable set of real numbers. If for any $x \in I$ we can find an integer $r(x)$ such that $f^{(r(x))}(x) \in A$, then such a function f is necessarily a polynomial. The interval I can be divided into three disjoint subsets: $S_1^{(f)}, S_2^{(f)}$, and $S_3^{(f)}$. These are characterized as follows: For $x \in S_1^{(f)}$ the formal Taylor series $\tau_x(f)$ of f around x converges to f in some neighborhood of x . For $x \in S_2^{(f)}$, $\tau_x(f)$ diverges. And for $x \in S_3^{(f)}$, $\tau_x(f)$ converges in some neighborhood of x but does not tend to f . Then $S_1^{(f)}$ is an open set and $S_2^{(f)}$ is a G_δ -set, while $S_3^{(f)}$ is an F_σ -set of the first category. Conversely, let $I = S_1 + S_2 + S_3$ be any partition of I into an open set S_1 , a G_δ -set S_2 , and an F_σ -set S_3 of the first category. Then there is some $f \in C^\infty(I)$ with $S_i = S_i^{(f)}$ ($i = 1, 2, 3$) [5].

E. Relations between C^∞ -Functions and Real Analytic Functions

Let $C^\omega(I)$ be the set of real analytic functions on I . Then $C^\omega(I)$ is a subalgebra of $C^\infty(I)$. Applying the above result in the case of $S_1 = \emptyset$, we find a function $f \in C^\infty(I)$ that admits no real analytic function coinciding with f in a subinterval of I . Actually, functions with such a property are distributed densely in $C^\infty(I)$. A necessary and sufficient condition for a function $f \in C^\infty(I)$ to belong to $C^\omega(I)$ is that for suitable constants A and k , $|f^{(n)}(x)| \leq Ak^n n!$, $x \in I, n = 0, 1, 2, \dots$, be valid (**Pringsheim's theorem**). If $f^{(n)}(x) \geq 0$ for all $x \in I$ and $n = 0, 1, 2, \dots$, then $f \in C^\omega(I)$ (S. N. Bernstein). For any open set $\Omega (\subset \mathbf{R}^n)$, the set $C^\omega(\Omega)$ of real analytic functions on Ω is dense in $C^\infty(\Omega)$ (**polynomial approximation theorem**). This result is true even when the topology of $C^\infty(\Omega)$

is replaced by a stronger one (Whitney [6]). Let $f \in C^\infty(\Omega)$ and $\varphi \in C^\infty(\Omega)$. Then we can find $g \in C^\infty(\Omega)$ satisfying $f = g\varphi$ if and only if for any $x \in \Omega$, $\tau_x(f)$ is divisible by $\tau_x(\varphi)$ in the ring of formal power series (S. Łojasiewicz, Malgrange [3]).

F. Quasi-Analytic Functions

The investigation of quasi-analytic functions began with the attempt to obtain an intrinsic characterization of analytic functions. Borel defined monogenic functions as functions differentiable on their domains of definition, which can be any subset of the complex plane, not necessarily assumed to be open (\rightarrow 198 Holomorphic Functions Q). Similar to complex analytic functions, monogenic functions are uniquely determined by their values on any curve. While quasi-analyticity can be defined by such properties, it is customary to approach quasi-analytic functions from another aspect, that is, the behavior of higher derivatives of C^∞ -functions.

Generally, a subset B of $C^\infty(I)$ is called a **set of quasi-analytic functions** if the mapping $\tau_x: B \rightarrow \mathbb{R}[[x]]$ defined in Section D is injective at each point $x \in I$. The functions belonging to B are called **quasi-analytic** (with respect to B). Here, an important problem is to characterize a set of quasi-analytic functions B by specializing the image of $\tau_x(B)$.

Now let $\{M_n\}$ be a sequence of positive numbers. Let $C(M_n)$ be the subset of $C^\infty(I)$ consisting of f such that

$$|f^{(n)}(x)| \leq A k^n M_n, \quad x \in I, \quad n = 0, 1, 2, \dots,$$

where $A = A(f)$ and $k = k(f)$ are constant. Then Pringsheim's theorem simply asserts that $C(n!) = C^\omega(I)$.

In 1912, J. Hadamard raised the problem of determining the condition that the sequence $\{M_n\}$ should satisfy so that $C(M_n)$ becomes a set of quasi-analytic functions [9]. A. Denjoy showed that if

$$M_n = (n \log^1 n \log^2 n \dots \log^p n)^n,$$

where

$$\log^1 n = \log n, \quad \log^p n = \log(\log^{p-1} n),$$

$$p = 2, 3, \dots,$$

then $C(M_n)$ is a set of quasi-analytic functions [10]. Later he derived an improved result that the condition $\sum M_n^{-1/n} = \infty$ is sufficient. T. Carleman first gave a necessary and sufficient condition for $C(M_n)$ to be a set of quasi-analytic functions, and later A. Ostrowski and T. Bang gave another version of the same condition [11–13]. The condition states essen-

tially the following: A necessary and sufficient condition for $C(M_n)$ to be a family of quasi-analytic functions in the interval (a, b) is given by either (i) $\sum \beta_n^{-1} = +\infty$, where $\beta_n = \inf_{k \geq n} M_k^{1/k}$ (Carleman), or (ii) $\int^\infty (\log T(r))/(r^2) dr = \infty$, where $T(r) = \sup_{n \geq 1} (r^n/M_n)$ (Ostrowski, Bang). S. Mandelbrojt and T. Bang also gave another condition [13, 14]. (The simplest proof of these results is found in [13] or [15], where the proof follows Bang's idea.)

Related to the above theorem, we also have the following: Let $\{M_n\}$ be a sequence of positive numbers with $\sum (M_n/M_{n+1}) < \infty$. For $\alpha > 0$ we can find $f \in C(M_n)$ defined on $(-\infty, \infty)$ such that $f(0) > 0$, $f^{(n)}(\pm\alpha) = 0$. Moreover, for $0 < \alpha < \beta$ there exists $f \in C(M_n)$ such that $f(0) > 0$, $f^{(n)} = 0$ ($\alpha \leq x \leq \beta$, $n = 0, 1, 2, \dots$) [16].

Suppose that we are given an interval I and increasing sequences $\{v_n\}$ and $\{M_n\}$ of positive numbers. Then we have the problem of finding suitable conditions on $\{v_n\}$ and $\{M_n\}$ under which the mapping $f \rightarrow \{f^{(v_n)}(x_0)\}$ gives an injective mapping from $C(M_n)$ to the sequences above. When $\{v_n\}$ and $\{M_n\}$ satisfy the above conditions, then a function belonging to $C(M_n)$ is called **quasi-analytic (v_n) in the generalized sense**. The study of the inclusion relation between two families $C(M_n)$ and $C(M'_n)$ also deserves attention. In [15] the relation between $C(M_n)$ and $C(n!) = C^\omega(I)$ is discussed in detail. There are many open problems concerning the relationship between $C(M_n)$ and $C(M'_n)$ in general.

Quasi-analytic functions are closely related to problems in various branches of analysis, in particular the theories of complex analytic functions, Fourier series, Fourier integrals, Dirichlet series, and asymptotic expansions [8, 15, 17].

G. Gevrey Class Functions

This class of functions has its origin in the study of \dagger parabolic equations. A C^∞ -function $f(x)$ defined in a domain of \mathbb{R}^n is called of **Gevrey class s** ($1 \leq s < \infty$) if for every compact set K in that domain there exist positive constants A_K and C_K such that

$$\max_K |D^\alpha f(x)| \leq A_K C_K^{|\alpha|} \alpha!^s$$

for all multi-indices α ($|\alpha| = \alpha_1 + \dots + \alpha_n$, $\alpha! = \alpha_1! \alpha_2! \dots \alpha_n!$). A typical example is

$$f(x) = \begin{cases} \exp(-1/x) & \text{for } x > 0, \\ 0 & \text{for } x \leq 0, \end{cases}$$

which is of Gevrey class 2. Let us denote the function space of a Gevrey class s by $\gamma^{(s)}$. Evidently $\gamma^{(s)} \subset \gamma^{(s')}$ if $s < s'$. Let $f(x) \in \gamma^{(s)}$ and $F(y) \in \gamma^{(s)}$; then $x \mapsto F(f(x)) \in \gamma^{(s)}$. In particular,

the sum and the product of two functions in $\gamma^{(s)}$ belong also to $\gamma^{(s)}$. Furthermore, the simplicity theorem holds in this class. Unlike the class of analytic functions, we can use this class of functions of class s (> 1) for a \dagger partition of unity, and several problems in partial differential equations can be treated within this function space. Gevrey class functions are discussed in a complete form in [20] (\rightarrow also [19]).

References

For functions of class C^∞ ,

- [1] R. P. Boas, A primer of real functions, Carus Math. Monographs, Wiley, 1960.
- [2] G. Glaeser, Fonctions composées différentiables, Ann. Math., (2) 77 (1963), 193–209.
- [3] B. Malgrange, Le théorème de préparation en géométrie différentiables, Sémin. H. Cartan, 1962–1963, Exp. 11, Inst. H. Poincaré, Univ. Paris, 1964.
- [4] B. Malgrange, Ideals of differentiable functions, Oxford Univ. Press, 1966.
- [5] H. Salzmann and K. Zeller, Singularitäten unendlich oft differenzierbarer Funktionen, Math. Z., 62 (1955), 354–367.
- [6] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, Trans. Amer. Math. Soc., 36 (1934), 63–89.
- [7] H. Whitney, Differentiable even functions, Duke Math. J., 10 (1943), 159–160.
- [8] H. Whitney, On ideals of differentiable functions, Amer. J. Math., 70 (1948), 635–658. For quasi-analytic functions,
- [9] J. Hadamard, Sur la généralisation de la notion de fonction analytique, Bull. Soc. Math. France, 40 (1912), 28–29.
- [10] A. Denjoy, Sur les fonctions quasi-analytiques de variable réelle, C. R. Acad. Sci. Paris, 173 (1921), 1329–1331.
- [11] T. Carleman, Les fonctions quasi-analytiques, Gauthier-Villars, 1926.
- [12] A. Ostrowski, Über quasianalytische Funktionen und Bestimmtheit asymptotischer Entwicklungen, Acta Math., 53 (1929), 181–266.
- [13] T. S. V. Bang, Om quasi-analytiske Funktioner, Thesis, Univ. of Copenhagen, 1946.
- [14] S. Mandelbrojt, Analytic functions and classes of infinitely differentiable functions, Rice Inst. Pamphlet 29, no. 1, 1942.
- [15] S. Mandelbrojt, Séries adhérentes, régularisation des suites, applications, Gauthier-Villars, 1952.
- [16] S. Mandelbrojt, Some theorems connected with the theory of infinitely differentiable functions, Duke Math. J., 11 (1949), 341–349.

- [17] S. Mandelbrojt, Séries de Fourier et classes quasi-analytiques de fonctions, Gauthier-Villars, 1935.
- [18] R. E. A. C. Paley and N. Wiener, Fourier transformations in the complex domain, Amer. Math. Soc. Colloq. Publ., 1934. For Gevrey class functions,
- [19] J. Hadamard, Les fonctions de classe supérieure dans l'équation de Volterra, J. Analyse Math., 1 (1951), 1–10.
- [20] M. Gevrey, Sur la nature des solutions des équations aux dérivées partielles, Ann. Sci. École Norm. Sup., ser. 3, 35 (1918), 243–304.

59 (V.14)

Class Field Theory

A. History

The notion of a class field was first introduced by D. Hilbert (1898). Let k be an \dagger algebraic number field and K a \dagger Galois extension of k . Hilbert called such a field K a class field over k (or K/k was called a class field) when the following property was satisfied: A \dagger prime ideal \mathfrak{p} of k of absolute degree 1 (i.e., a prime ideal whose \dagger absolute norm is a prime number) is decomposed in K as the product of prime ideals of K of absolute degree 1 if and only if \mathfrak{p} is a \dagger principal ideal. (Such a field K is now said to be an **absolute class field** over k in order to distinguish it from a class field defined later more generally by T. Takagi as explained below.) Hilbert conjectured the following theorems (1)–(4) together with the principal ideal theorem (\rightarrow Section D), and proved them in some special cases. (1) For any algebraic number field k there exists one and only one class field K/k . (2) A class field K/k is an \dagger Abelian extension whose \dagger Galois group is isomorphic to the \dagger ideal class group of k . Hence the degree $n = [K:k]$ is equal to the \dagger class number h of k . (3) The \dagger relative different of a class field K/k is the principal order; thus K/k is an \dagger unramified extension. (4) Let \mathfrak{p} be a prime ideal of k , and let f be the smallest positive integer such that \mathfrak{p}^f is a principal ideal. Then \mathfrak{p} is decomposed in the class field K/k as $\mathfrak{p} = \mathfrak{P}_1 \mathfrak{P}_2 \dots \mathfrak{P}_g$, $N_{K/k}(\mathfrak{P}_i) = \mathfrak{p}^f$, $fg = n$.

Hilbert was led to these conjectures by the analog to the theory of \dagger algebraic functions in one variable. Theorems (1)–(4) were proved by P. Furtwängler (*Math. Ann.*, 63 (1907)), but these results were subsumed under the class field theory of Takagi, who generalized the notion of class field and proved that every Abelian extension of k is a class field over k (*J. Coll. Sci. Imp. Univ. Tokyo*, (9) 41 (1920)). Since

then, the arithmetic of Abelian extensions of k has developed through this theory. In Takagi's paper, L. Kronecker's problem concerning Abelian extensions of an imaginary quadratic field (\rightarrow 73 Complex Multiplication) was solved simultaneously; this had been an open problem since the 19th century. Later, E. Artin proved the general law of reciprocity (*Abh. Math. Sem. Univ. Hamburg*, 5 (1927)) which put class field theory into its complete form. The original proof by Takagi was rather complicated, and H. Hasse, Artin, J. Herbrand, C. Chevalley, and others tried to simplify it. In particular, Chevalley introduced the notion of \dagger ideles and gave a purely arithmetic proof. On the other hand, attempts are also being made to generalize this theory to non-Abelian extensions. We mention here the results of G. Shimura [15] and Y. Ihara [16]. Today, class field theory is considered one of the most beautiful theories in mathematics.

B. Definition of a Class Field

Let k be an algebraic number field. For the definition of a general class field over k , we need a generalization of the ideal class group of k (\rightarrow 14 Algebraic Number Fields H). Let \mathfrak{m} be an \dagger integral divisor of k , and let $\mathfrak{I}(\mathfrak{m})$ be the multiplicative group of all \dagger fractional ideals of k which are \dagger relatively prime to \mathfrak{m} . For the rest of this article, we mean by an ideal of k a fractional ideal of k . Denote by $S(\mathfrak{m})$ the \dagger ray modulo \mathfrak{m} . Let $H(\mathfrak{m})$ be an \dagger ideal group modulo \mathfrak{m} , that is, a subgroup of $\mathfrak{I}(\mathfrak{m})$ containing $S(\mathfrak{m})$. A Galois extension K of k is said to be a **class field** over k for the ideal group $H(\mathfrak{m})$ if the following property is satisfied: A prime ideal \mathfrak{p} of k of absolute degree 1 which is relatively prime to \mathfrak{m} is decomposed in K as the product of prime ideals of K of absolute degree 1 if and only if \mathfrak{p} belongs to $H(\mathfrak{m})$. The absolute class field of Hilbert is the case where $\mathfrak{m}=(1)$ and $H(\mathfrak{m})$ is the group of all principal ideals of k .

A class field K/k for an ideal group H is uniquely determined by H (**uniqueness theorem**). The \dagger conductor \mathfrak{f} of H is said to be the **conductor** of the class field for H . The ideal group H corresponding to the class field K/k is determined by K as follows: $H(\mathfrak{f})/S(\mathfrak{f})$ is the union of all cosets C of $\mathfrak{I}(\mathfrak{f})$ modulo $S(\mathfrak{f})$ such that C contains a \dagger relative norm $N_{K/k}(\mathfrak{A})$ of some ideal \mathfrak{A} of K which is relatively prime to \mathfrak{f} . In general, let K/k be a Galois extension and \mathfrak{m} be an integral divisor of k . Let $H(\mathfrak{m})$ be the union of all cosets C of $\mathfrak{I}(\mathfrak{m})$ modulo $S(\mathfrak{m})$ such that C contains a relative norm $N_{K/k}(\mathfrak{A})$ of some ideal \mathfrak{A} of K which is relatively prime to

\mathfrak{m} . Then $H(\mathfrak{m})$ is a multiplicative subgroup of $\mathfrak{I}(\mathfrak{m})$, and the index $h=(\mathfrak{I}(\mathfrak{m}):H(\mathfrak{m}))$ is not greater than the degree $n=[K:k]$. We have $h=n$ if and only if K/k is the class field for H . Hence a class field K/k can be defined as a Galois extension of k such that $h=n$ for a suitable integral divisor \mathfrak{m} of k .

C. Fundamental Theorems in Class Field Theory

(1) **Main theorem**: Any Abelian extension K/k is a class field over k for a suitable ideal group H .

(2) **Existence theorem**: For any ideal group $H(\mathfrak{m})$ there exists one and only one class field for $H(\mathfrak{m})$.

(3) **Composition theorem**: Let K_1 and K_2 be class fields for H_1 and H_2 , respectively. Then the composite field $K_1 K_2$ is the class field over k for $H_1 \cap H_2$. Consequently, $K_1 \supset K_2$ if and only if $H_1 \subset H_2$.

(4) **Isomorphism theorem**: The Galois group of a class field K/k for $H(\mathfrak{m})$ is isomorphic to $\mathfrak{I}(\mathfrak{m})/H(\mathfrak{m})$. In particular, every class field K/k is an Abelian extension of k .

(5) **Decomposition theorem**: Let \mathfrak{f} be the conductor of the class field for H . If \mathfrak{p} is a prime ideal of k relatively prime to \mathfrak{f} and f is the smallest positive integer with $\mathfrak{p}^f \in H$, then \mathfrak{p} is decomposed in K/k as $\mathfrak{p} = \mathfrak{P}_1 \mathfrak{P}_2 \dots \mathfrak{P}_g$, $N_{K/k}(\mathfrak{P}_i) = \mathfrak{p}^f$, $fg = n$.

(6) **Conductor-ramification theorem**: Let \mathfrak{f} be the conductor of a class field K/k . Then \mathfrak{f} is not divisible by any prime divisor that is unramified for K/k , and \mathfrak{f} is divisible by every prime divisor that ramifies for K/k . Let $\mathfrak{f} = \prod \mathfrak{f}_p$, $\mathfrak{f}_p = \mathfrak{p}^c$. Then \mathfrak{f}_p coincides with the \dagger \mathfrak{p} -conductor of K/k , and the exponent c can be explicitly expressed by the order of the \dagger ramification groups and the \dagger ramification numbers of \mathfrak{p} for K/k (\rightarrow 14 Algebraic Number Fields P).

(7) Let \mathfrak{p} be a prime ideal of k that ramifies for K/k . Let H_p be the ideal group of k such that (i) the conductor of H_p is relatively prime to \mathfrak{p} and (ii) H_p is the minimal ideal group of k containing H with property (i). Let $n=[K:k]$, $e=(H_p:H)$, and $\mathfrak{p}^f \in H_p$, where $\mathfrak{p}^d (d < f) \notin H_p$. Then \mathfrak{p} is decomposed in K/k as $\mathfrak{p} = (\mathfrak{P}_1 \mathfrak{P}_2 \dots \mathfrak{P}_g)^e$, $N_{K/k}(\mathfrak{P}_i) = \mathfrak{p}^f$, $n = efg$.

(8) **Translation theorem**: Let K/k be the class field for an ideal group $H(\mathfrak{m})$, and let Ω be an arbitrary finite extension of k . Then $K\Omega/\Omega$ is the class field for H^* , where H^* is the ideal group of Ω consisting of all ideals \mathfrak{b} of Ω with $N_{\Omega/k}(\mathfrak{b}) \in H(\mathfrak{m})$. In particular, the conductor of $K\Omega/\Omega$ is a divisor of the conductor of K/k .

(9) **Artin's general law of reciprocity**: Let K/k be the class field for an ideal group H with the

conductor \mathfrak{f} . We denote the * Artin symbol of an ideal \mathfrak{a} of k that is relatively prime to \mathfrak{f} by

$$(K/\mathfrak{a}) = \left(\frac{K/k}{\mathfrak{a}} \right).$$

Let a mapping Φ from $\mathfrak{I}(\mathfrak{f})$ to the Galois group G of K/k be defined by $\Phi(\mathfrak{a}) = (K/\mathfrak{a})$ for $\mathfrak{a} \in \mathfrak{I}(\mathfrak{f})$. Then Φ induces the isomorphism $\mathfrak{I}(\mathfrak{f})/H(\mathfrak{f}) \cong G$. Namely, the isomorphism mentioned in (4) is explicitly given by the Artin symbol. Also, the ideal group $H(\mathfrak{f})$ is characterized as the set of all ideals \mathfrak{a} such that $\mathfrak{a} \in \mathfrak{I}(\mathfrak{f})$ and $(K/\mathfrak{a}) = 1$. From this theorem we can prove all the known laws of reciprocity for power-residue and norm-residue symbols (\rightarrow 14 Algebraic Number Fields O, Q, R).

From the general results of class field theory we can systematically derive all the known theorems concerning the arithmetic of quadratic fields, cyclotomic fields, and Kummer extensions.

D. Principal Ideal Theorem

Let K/k be an absolute class field. Then the extension of any ideal of k to K is a principal ideal of K . This theorem is called the **principal ideal theorem**. It was conjectured by Hilbert, formulated by Artin as a theorem of group theory, and proved by Furtwängler (*Abh. Math. Sem. Univ. Hamburg*, 7 (1930)). Later a simple proof was given by S. Iyanaga (*Abh. Math. Sem. Univ. Hamburg*, 10 (1934)). This theorem was also generalized to the following general principal ideal theorem (Iyanaga, *Japan. J. Math.*, 7 (1930)): Let K/k be the class field for the ray $S(\mathfrak{f})$, and let $\mathfrak{f} = \mathfrak{f}\mathfrak{D}$, where \mathfrak{D} is the relative different of K/k . Then the extension to K of any ideal of k that is relatively prime to \mathfrak{f} belongs to $S(\mathfrak{f})$. Put $\mathfrak{F} = \prod \mathfrak{P}^v$. Then v is equal to the ramification number $v_r + 1$ (\rightarrow 14 Algebraic Number Fields K). For an absolute class field K/k , let the extension of an ideal \mathfrak{a} to K be $(\Theta(\mathfrak{a}))$. Then we can choose $\Theta(\mathfrak{a}) \in K$ such that $\Theta(\mathfrak{a})\Theta(\mathfrak{b})^{\sigma(\mathfrak{a})}\Theta(\mathfrak{a}\mathfrak{b})^{-1} \in k$, where $\sigma(\mathfrak{a}) = (K/\mathfrak{a})$ is the Artin symbol for \mathfrak{a} (T. Tannaka, *Ann. Math.*, 67 (1958)). This result can also be generalized for the class field for $S(\mathfrak{f})$.

E. Theory of Genera

Let K/k be a Galois extension and let $H(\mathfrak{m})$ be an ideal group of k . The set of all ideals \mathfrak{A} of K relatively prime to \mathfrak{m} such that $N_{K/k}(\mathfrak{A})$ belongs to $H(\mathfrak{m})$ forms an ideal group of K . This ideal group is said to be the **principal genus** for H . Each coset of $\mathfrak{I}(\mathfrak{m})$ modulo the principal genus for H is said to be a **genus** for H . In

particular, let K/k be a cyclic extension with the conductor \mathfrak{f} , and let $H(\mathfrak{f})$ be the ideal group of k generated by $N_{K/k}(A)$ ($A \in K$) and $S(\mathfrak{f})$. Then the principal genus for $H(\mathfrak{f})$ is the ideal group formed by the ideal classes of K of the form $C^{1-\sigma}$, where σ is a generator of the Galois group of K/k (\rightarrow 347 Quadratic Fields F). In general, let K/k be an Abelian extension, and let \mathfrak{f} be the conductor of K/k . Then for the ideal $\mathfrak{F} = \prod \mathfrak{P}^v$ of K defined in Section D, $N_{K/k}(S(\mathfrak{m}\mathfrak{F})) = S(\mathfrak{m}\mathfrak{f})$ for an arbitrary integral ideal \mathfrak{m} of k . In particular, let K/k be a cyclic extension, and let $H = S(\mathfrak{m}\mathfrak{f})$. Then the principal genus for H is the ideal group consisting of all cosets of $\mathfrak{I}(\mathfrak{m}\mathfrak{F})$ modulo $S(\mathfrak{m}\mathfrak{F})$ of the form $B^{1-\sigma}$ (Herbrand; Iyanaga, *J. Reine Angew. Math.*, 171 (1934)).

F. Class Field Tower Problem and Construction Problem

Furtwängler considered the following problem: Let k be a given algebraic number field, $k = k_0 \subset k_1 \subset k_2 \dots$ be the sequence of fields such that k_i is the absolute class field over k_{i-1} , and K_∞ be the union of all the k_i . Is K_∞ a finite extension of k ? The answer is yes if and only if k_n is of class number 1 for some n . This problem is called the **class field tower problem**. Artin remarked that if for every algebraic number field F of degree n we have the inequality $|D_F| > (\pi/4)^{2r_2} (n^n/n!)^2 > (\pi e^2/4)^n / (2\pi n e^{1/6n})$ for the * discriminant D_F , then K_∞/k is always finite [1, p. 46]. E. S. Golod and I. R. Shafarevich (1964) solved the class field tower problem negatively; they proved that K_∞/k is infinite if k_i ($i = 1, 2, \dots$) is the maximal unramified Abelian p -extension of k_{i-1} for a fixed prime number p and if the inequality $\gamma \geq 3 + 2\sqrt{\rho + 2}$ holds, where γ is the minimal number of generators of the p -component of the ideal class group of k and ρ is the rank of the unit group of k . (We call an extension K/k a **p -extension** if the degree $[K:k]$ is a power of a prime number p .) For example, the class field tower K_∞/k is actually infinite if k is an imaginary quadratic field ($\rho = 1$) and $\gamma \geq 7$ for $p = 2$, for example, $k = Q(\sqrt{-3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19})$.

Construction problem. Let k be a given algebraic number field and G a finite group. The construction problem asks us whether there exists a Galois extension K/k such that its Galois group $\text{Gal}(K/k)$ is isomorphic to G . If G is Abelian the problem can be solved affirmatively by using class field theory. This problem was also solved affirmatively for p -groups by A. Scholtz and H. Reinhardt in 1937, and for general solvable groups by Shafarevich in 1954 (*Izv. Akad. Nauk SSSR*, ser. mat. 18).

G. Class Field Theory for Algebraic Function Fields and Local Class Field Theory

F. K. Schmidt developed an analog of class field theory for Abelian extensions over an algebraic function field in one variable with finite coefficient field (1930; → [8]). An arithmetic proof was given by M. Moriya (1938). An analog of class field theory for local fields with finite residue-class fields, called **local class field theory** (→ 257 Local Fields) was first developed by Hasse, and later Chevalley gave an algebraic derivation (1933).

H. Cohomology of Groups and Class Field Theory

For the purpose of simplifying the proof of the main theorems in class field theory, the theory of \dagger Galois cohomology was developed by T. Nakayama, G. Hochschild, A. Weil, Artin, J. Tate, and others. In particular, Artin and Tate [9] constructed class field theory on the basis of the cohomology theory of finite groups as follows: Let G be a finite group, A a $\dagger G$ -module (or a multiplicative commutative group with the operator domain G), and $\hat{H}^n(G, A)$ the n th \dagger cohomology group ($n=0, \pm 1, \pm 2, \dots$) of G with coefficients in A (→ 200 Homological Algebra N). Then we have $\hat{H}^0(G, A) \cong A^G/N_G(A)$, where A^G is the set of all G -invariant elements in A and $N_G(A)$ is the set of all elements of the form $N_G(a) = \sum_{\sigma \in G} \sigma a$ ($a \in A$). We can consider \mathbf{Z} a G -module by defining $\sigma n = n$ ($n \in \mathbf{Z}, \sigma \in G$). Let A, B, C be G -modules such that a G -bilinear mapping $(A, B) \rightarrow C$ is defined. Then we can define the \dagger cup product $(\alpha, \beta) \rightarrow \alpha \smile \beta$ ($\alpha \in \hat{H}^r(G, A), \beta \in \hat{H}^s(G, B), \alpha \smile \beta \in \hat{H}^{r+s}(G, C)$) for $r, s \in \mathbf{Z}$ with the usual properties. Let A be a G -module and H a subgroup of G . Then the \dagger restriction homomorphism $R_{G/H}: \hat{H}^n(G, A) \rightarrow \hat{H}^n(H, A)$ and the \dagger injection homomorphism $\text{Inj}_{H/G}: \hat{H}^n(H, A) \rightarrow \hat{H}^n(G, A)$ are defined for $n \in \mathbf{Z}$. If H is a normal subgroup of G , then the \dagger inflation homomorphism $\text{Inf}_{(G/H)/G}: \hat{H}^n(G/H, A^H) \rightarrow \hat{H}^n(G, A)$ can be defined for $n \geq 1$ (→ 200 Homological Algebra M).

Let k be an algebraic number field, and let K be a Galois extension of k of degree n with the Galois group $G = G(K/k)$. The multiplicative group $K^\times = K - \{0\}$, the \dagger idele group J_K of K , and the idele class group C_K of K are multiplicative commutative groups with G as their operator domain. The fundamental formulas in Galois cohomology for class field theory are

$$\hat{H}^1(G, C_K) = 0, \quad (1)$$

$$\hat{H}^2(G, C_K) \cong \mathbf{Z}/n\mathbf{Z}. \quad (2)$$

It is possible to realize the isomorphism of (2) by the **invariant** $\text{inv}_{K/k}: \hat{H}^2(G, C_K) \xrightarrow{\sim} \{(r/n) \pmod{\mathbf{Z}} \mid r=0, 1, \dots, n-1\}$ in such a way that the following properties hold, where the **canonical cohomology class** for K/k is the element $\xi_{K/k}$ of $\hat{H}^2(G, C_K)$ such that $\text{inv}_{K/k} \xi_{K/k} = (1/n) \pmod{\mathbf{Z}}$: (i) for $k \subset l \subset K, G = G(K/k), H = G(K/l)$, the relation $\text{Res}_{G/H} \xi_{K/k} = \xi_{K/l}$ holds; (ii) if l/k is also a Galois extension with $F = G(l/k)$, then we have $\text{Inf}_{F/G} \xi_{l/k} = \xi_{K/k}^m$ ($m = [K:l]$); (iii) for a cyclic extension K/k we have $\text{inv}_{K/k} \xi_{K/k} = \sum_{\mathfrak{p}} \text{inv}_{\mathfrak{p}}(\xi_{K/k})_{\mathfrak{p}} \pmod{\mathbf{Z}}$, where \mathfrak{p} runs over all prime divisors of k and $\text{inv}_{\mathfrak{p}}$ is the invariant in the local theory (→ 257 Local Fields E). By these properties, the canonical cohomology class $\xi_{K/k}$ is uniquely determined. After these preliminaries we can state Tate's theorem, from which the fundamental theorems in class field theory follow.

Tate's theorem. Let K/k be a Galois extension with the Galois group G . Then we have the isomorphism $\Phi_n: \hat{H}^{n-2}(G, \mathbf{Z}) \cong \hat{H}^n(G, C_K)$ ($n = 0, \pm 1, \pm 2, \dots$) that is given explicitly by $\Phi_n(\alpha) = \xi_{K/k} \smile \alpha$, where $\xi_{K/k} \in \hat{H}^2(G, C_K)$ is the canonical cohomology class for K/k (*Ann. Math.*, (2) 56 (1952)).

Corollary 1. Since $\hat{H}^{-2}(G, \mathbf{Z}) \cong G/[G, G]$ and $\hat{H}^0(G, C_K) \cong C_k/N_{K/k}(C_K)$, we have the isomorphism $\Phi_0: G/[G, G] \cong C_k/N_{K/k}(C_K)$. Let $f(\tau, \sigma)$ ($\tau, \sigma \in G$) be a 2- \dagger cocycle belonging to $\xi_{K/k}$. Then by the explicit expression for the cup product we obtain the isomorphism

$$\Phi_0: \sigma \pmod{[G, G]} \rightarrow \prod_{\tau \in G} f(\tau, \sigma)^{-1} \pmod{N_{K/k}(C_K)}.$$

This is an analog of the result in local theory that was proved earlier by T. Nakayama and Y. Akizuki (*Math. Ann.*, 112 (1936)).

Corollary 2. For an Abelian extension K/k we have the isomorphism $\Phi_0: G \cong C_k/N_{K/k}(C_K)$. Φ_0^{-1} has the property of being the norm-residue symbol for C_k , and from this isomorphism we can prove immediately Artin's law of reciprocity. Thus we can prove the main theorems in class field theory by cohomology-theoretic methods [9].

We can also see, by generalizing this isomorphism to infinite Abelian extensions, that the Galois group of the maximal Abelian extension of k over the ground field k with \dagger Krull topology is algebraically and topologically isomorphic to C_k/D_k , where D_k is the connected component of the unit element in C_k . The structures of D_k and C_k/D_k were explicitly determined by Artin and T. Kubota, respectively (→ 6 Adeles and Ideles D).

If we assume the fundamental formulas (1) and (2) stated above and several other simple assumptions as axioms for an infinite ex-

tension of a fixed ground field, we can develop the results stated in this section purely cohomology-theoretically. Such a system is called a **class formation** (Artin; \rightarrow [9]). In addition to the cases of algebraic number fields, algebraic function fields in one variable with finite coefficient fields, and local fields with finite residue-class fields, which we have mentioned already, we also know several other cases for which analogies of class field theory are valid. These analogies can be explained systematically by using class formation theory (Y. Kawada, *Duke Math. J.*, 22 (1955)). Examples are (1) the theory of unramified Abelian extensions of an algebraic function field in one variable with algebraically closed constant field of characteristic 0 (Tate and Kawada, *Amer. J. Math.*, 77 (1955)); (2) the theory of Kummer extensions over a field k such that (i) the characteristic of k is 0, (ii) k contains all the roots of unity, and (iii) for any Galois extension K/k , $N_{K/k}(K) = k$; (3) the theory of Abelian p -extensions of a field of characteristic p (E. Witt, *J. Reine Angew. Math.*, 176 (1963); I. Satake and Kawada, *J. Fac. Sci. Univ. Tokyo*, 7 (1955)); (4) the theory of unramified Abelian p -extensions of an algebraic function field in one variable with algebraically closed constant field of characteristic p (Hasse and Witt, *Monatsh. Math.*, 43 (1936), H. L. Schmid, I. R. Shafarevich, Kawada, T. Tamagawa); and (5) the theory of Abelian extensions of a local field with algebraically closed residue-class fields (J.-P. Serre, *Bull. Soc. Math. France*, 89 (1961)).

An analogy of class field theory for infinite Abelian extensions was considered by Herbrand, Moriya, M. Mori, and Kawada.

References

- [1] H. Hasse, Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper, Jber. Deutsch. Math. Verein. I, 35 (1926), 1–55; Ia, 36 (1927), 231–311; II, 1930 (Physica Verlag, 1965).
- [2] H. Hasse, Vorlesungen über Klassenkörpertheorie, Lecture notes, Univ. Marburg, 1932 (Physica Verlag, 1967).
- [3] C. Chevalley, Sur la théorie du corps des classes dans les corps finis et les corps locaux, J. Fac. Sci. Univ. Tokyo, 2 (1933), 365–476.
- [4] J. Herbrand, Le développement moderne de la théorie des corps algébriques, Gauthier-Villars, 1936.
- [5] E. Artin, Algebraic numbers and algebraic functions, Lecture notes, Princeton Univ., 1950–1951 (Gordon & Breach, 1967).
- [6] C. Chevalley, La théorie du corps de classes, Ann. Math., (2) 41 (1940), 394–418.

- [7] A. Weil, Sur la théorie du corps de classes, J. Math. Soc. Japan, 3 (1951), 1–35.
- [8] F. K. Schmidt, Die Theorie des Klassenkörpers über einem Körper algebraischer Funktionen in einer Unbestimmten und mit endlichem Konstantenbereich, S.-B. Phys.-Med. Soz. Erlangen, 62 (1931), 267–284.
- [9] E. Artin and J. Tate, Class field theory, Princeton Univ. Press, 1951 (Benjamin, 1967).
- [10] S. Iyanaga, Class field theory notes, Univ. of Chicago, 1961.
- [11] J.-P. Serre, Groupes algébriques et corps de classes, Actualités Sci. Ind., Hermann, 1959.
- [12] J.-P. Serre, Corps locaux, Actualités Sci. Ind., Hermann, 1962; English translation, Local fields, Springer, 1979.
- [13] A. Weil, Basic number theory, Springer, 1967.
- [14] J. W. S. Cassels and A. Fröhlich (eds.), Algebraic number theory, Academic Press, 1967.
- [15] G. Shimura, A reciprocity law in nonsolvable extensions, J. Reine Angew. Math., 221 (1966), 209–220.
- [16] Y. Ihara, Non-Abelian class fields over function fields in special cases, Actes Congr. Intern. Math., 1970, Nice, Gauthier-Villars, 381–389.
- [17] Y. Ihara, Some fundamental groups in the arithmetic of algebraic curves over finite fields, Proc. Nat. Acad. Sci. US, 72 (1975), 3281–3284.
- [18] Y. Kawada, Class formations, 1969 Number Theory Institute, Amer. Math. Soc. Proc. Symposia in Pure Math., 20 (1971), 96–114.
- [19] S. Iyanaga (ed.), Theory of numbers, North-Holland, 1975. (Original in Japanese, 1969.)

60 (IV.5) Classical Groups

A. Introduction

The general linear groups, unitary groups, orthogonal groups, symplectic groups, etc., that are described below are all called **classical groups** (\rightarrow 13 Algebraic Groups, 151 Finite Groups, 248 Lie Algebras, 249 Lie Groups).

B. General Linear Groups

Let V be a \dagger linear space of dimension n over a \dagger field K , and let $GL(V)$ denote the set of all \dagger linear mappings of V onto V (hence they are

all bijections). Then $GL(V)$ is a group under the composition of mappings. This group is called the **general linear group** (or **full linear group**) on V . Let e_1, \dots, e_n be a basis of V over K , and let (α_j^i) be the matrix associated with an element A of $GL(V)$: $Ae_i = \sum_j \alpha_j^i e_j$. Then the mapping $A \rightarrow (\alpha_j^i)$ is an isomorphism of $GL(V)$ onto the multiplicative group $GL(n, K)$ of all $n \times n$ invertible matrices over K . We can thus identify the group $GL(V)$ with $GL(n, K)$. $GL(n, K)$ is called the **general linear group of degree n over K** . Consider the homomorphism $A \rightarrow |A|$ ($|A|$ is the determinant of A) of $GL(V)$ onto the multiplicative group $K^* = K - \{0\}$. Its kernel $SL(V)$ is a normal subgroup of $GL(V)$ and is called the **special linear group** (or **unimodular group**) on V . The subgroup $SL(n, K) = \{A \mid A \in GL(n, K), |A| = 1\}$ of $GL(n, K)$ corresponds to $SL(V)$ under the above isomorphism $GL(V) \cong GL(n, K)$. $SL(n, K)$ is called the **special linear group of degree n over K** . Unless $n = 2$ and K is the finite field $F_2 = GF(2)$, $SL(n, K)$ is the commutator subgroup of $GL(n, K)$. The center \mathfrak{z} of $GL(n, K)$ coincides with the set of all scalar matrices αI ($\alpha \in K^*$), and the center \mathfrak{z}_0 of $SL(n, K)$ is a finite group given by $\mathfrak{z} \cap SL(n, K) = \{\alpha I \mid \alpha \in K, \alpha^n = 1\}$.

Now let $P(V)$ be the projective space of dimension $n - 1$ obtained from a linear space V of dimension n . Namely, $P(V)$ is the set of all linear subspaces of dimension 1. Then there exists a natural homomorphism ϕ of $GL(V)$ into the group of all projective transformations of $P(V)$, and the kernel of ϕ coincides with the center \mathfrak{z} of $GL(V)$. Hence $\phi(GL(V)) \cong GL(V)/\mathfrak{z}$. This group is written as $PGL(V)$ and is called the **projective general linear group** on $P(V)$. Similarly, $PGL(n, K) = GL(n, K)/\mathfrak{z}$ is called the **projective general linear group of degree n over K** . The quotient group $SL(n, K)/\mathfrak{z}_0$ of $SL(n, K)$ by the center \mathfrak{z}_0 is called the **projective special linear group** and is written as $PSL(n, K)$ or $LF(n, K)$ (**linear fractional group**).

The groups $GL(n, K)$, $SL(n, K)$, etc., are also written as $GL_n(K)$, $SL_n(K)$, etc. In particular, when K is the finite field F_q , these groups are denoted by $GL(n, q)$, $SL(n, q)$, $PGL(n, q)$, $PSL(n, q)$, $LF(n, q)$.

Simplicity of $PSL(n, k)$. When $n = 2$ and $K = F_2$, $PSL(2, 2) \cong \mathfrak{S}_3$ (the symmetric group of degree 3). When $n = 2$ and $K = F_3$, $PSL(2, 3) \cong \mathfrak{A}_4$ (the alternating group of degree 4). Except for these cases, the group $PSL(n, K)$ ($n \geq 2$) is a noncommutative simple group (\rightarrow 151 Finite Groups I).

Suppose that K is the finite field F_q , and let $\alpha(n, q)$, $\beta(n, q)$, $\gamma(n, q)$, $\delta(n, q)$ denote the orders of $GL(n, q)$, $SL(n, q)$, $PGL(n, q)$, $PSL(n, q)$, res-

pectively. Then we have

$$\alpha(n, q) = (q^n - 1)(q^n - q) \dots (q^n - q^{n-1}),$$

$$\beta(n, q) = \gamma(n, q) = \alpha(n, q)/(q - 1),$$

$$\delta(n, q) = \gamma(n, q)/d,$$

where $d = (n, q - 1)$ (the greatest common divisor of n and $q - 1$).

C. Properties as Lie Groups

If the ground field K is the field \mathbf{R} of real numbers (the field \mathbf{C} of complex numbers), the above groups are all Lie groups (complex Lie groups). In particular, $SL(n, \mathbf{C})$ is a simply connected, simple, and semisimple complex Lie group of type A_{n-1} , and $PSL(n, \mathbf{C})$ is the adjoint group of the complex simple Lie algebra of type A_{n-1} .

D. Determination of the Rational Representations of $GL(V)$

In Sections D and E, the field K is assumed to be of characteristic 0. Let ρ be a homomorphism of $GL(V) = GL(n, K)$ into $GL(m, K)$ ($\rho: A = (\alpha_j^i) \rightarrow B = (\beta_j^i)$). Then if each β_j^i is a rational function (or polynomial or analytic function) in $(\alpha_1^1, \alpha_2^1, \dots, \alpha_n^n)$ over K , ρ is called a **rational representation** (or **polynomial** or **analytic representation**) of degree m of $GL(V)$. (We suppose that K is \mathbf{R} or \mathbf{C} when we consider analytic functions.) For example, every rational representation of degree 1 can be expressed as $A \rightarrow |A|^e$ (e is an integer). In particular, if K is the field \mathbf{C} of complex numbers, every analytic representation of $GL(n, \mathbf{C})$ is a rational representation. Since $GL(n, \mathbf{C})$ is the complexification of the unitary group $U(n)$, there exists a one-to-one correspondence between the complex analytic representations of $GL(n, \mathbf{C})$ and the continuous representations of $U(n)$; this correspondence preserves equivalence, irreducibility, tensor product, and direct sum of the representations (\rightarrow 249 Lie Groups). Hence, determining the rational representations of $GL(n, \mathbf{C})$ is equivalent to determining the continuous representations of $U(n)$. In the general case, the rational representations of $GL(V) = GL(n, K)$ are all completely reducible. For any rational representation ρ of $GL(V)$, there exists a natural number e such that the representation $\rho': A \rightarrow |A|^e \rho(A)$ is a polynomial representation. Hence in order to determine the rational representations of $GL(V)$, it is sufficient to determine the irreducible polynomial representations of $GL(V)$, which, as described below, can be obtained by decomposing the representations on the tensor space

$V^m = V \otimes \dots \otimes V$ of degree m (m copies of V) ($m = 1, 2, \dots$). For $A \in GL(V)$, define $D_m(A) \in GL(V^m)$ as the tensor product

$$D_m(A) = A \otimes \dots \otimes A \text{ (} m \text{ copies of } A \text{)}.$$

Namely, for $v_1, \dots, v_m \in V$, we have

$$D_m(A)(v_1 \otimes \dots \otimes v_m) = Av_1 \otimes \dots \otimes Av_m.$$

The mapping $GL(V) \ni A \rightarrow D_m(A) \in GL(V^m)$ is a polynomial representation of degree n^m of $GL(V)$. Now let $\mathfrak{Q}(V^m)$ be the \dagger associative algebra of all linear mappings of V^m into V^m (\dagger total matrix algebra), and let \mathfrak{A} be the subalgebra of $\mathfrak{Q}(V^m)$ generated by $\{D_m(A) \mid A \in GL(V)\}$. Next, for an element σ of the symmetric group \mathfrak{S}_m of degree m , define $B_\sigma \in GL(V^m)$ by $B_\sigma(v_1 \otimes \dots \otimes v_m) = v_{\sigma^{-1}(1)} \otimes \dots \otimes v_{\sigma^{-1}(m)}$. Then the mapping $\sigma \rightarrow B_\sigma$ is a representation of \mathfrak{S}_m on V^m . Thus we obtain a representation τ of the \dagger group ring $K[\mathfrak{S}_m]$ of \mathfrak{S}_m over K on $V^m: K[\mathfrak{S}_m] \rightarrow \mathfrak{Q}(V^m)$. Set $\tau(K[\mathfrak{S}_m]) = \mathfrak{B}$. Then \mathfrak{A} and \mathfrak{B} are \dagger commutators of each other in $\mathfrak{Q}(V^m)$, i.e., $\mathfrak{A} = \{X \in \mathfrak{Q}(V^m) \mid XB = BX \text{ (for all } B \in \mathfrak{B})\}$, $\mathfrak{B} = \{X \in \mathfrak{Q}(V^m) \mid AX = XA \text{ (for all } A \in \mathfrak{A})\}$.

Now for a right ideal \mathfrak{r} of \mathfrak{B} , let $\mathfrak{r}(V^m)$ be the subspace of V^m composed of all the finite sums of the form $\sum Bx$ ($B \in \mathfrak{r}, x \in V^m$). Then the following statements hold:

(1) $\mathfrak{r}(V^m)$ is invariant under \mathfrak{A} ; hence it is a subspace of V^m invariant under $GL(V)$. Conversely, for any subspace U of V^m invariant under $GL(V)$, there exists a unique right ideal \mathfrak{r} of \mathfrak{B} such that $U = \mathfrak{r}(V^m)$.

(2) Let $\mathfrak{r}_1, \mathfrak{r}_2$ be right ideals of \mathfrak{B} , and put $U_1 = \mathfrak{r}_1(V^m)$, $U_2 = \mathfrak{r}_2(V^m)$. Then $\mathfrak{r}_1 \cong \mathfrak{r}_2$ (as right \mathfrak{S}_m -modules) if and only if $U_1 \cong U_2$ (as representation spaces of $GL(V)$).

(3) The mapping $\mathfrak{r} \rightarrow \mathfrak{r}(V^m)$ is a lattice isomorphism of the \dagger lattice of right ideals of \mathfrak{B} onto the lattice of $GL(V)$ -invariant subspaces of V^m . Hence if $\mathfrak{r} = \mathfrak{r}_1 + \mathfrak{r}_2$ (direct sum), then $U = U_1 + U_2$ (direct sum). Also, $\mathfrak{r}(V^m)$ gives an irreducible representation of $GL(V)$ if and only if \mathfrak{r} is a minimal right ideal of \mathfrak{B} .

Since the algebra $K[\mathfrak{S}_m]$ is a \dagger semisimple algebra, \mathfrak{B} can be considered as a two-sided ideal of $K[\mathfrak{S}_m]$. Hence a minimal right ideal \mathfrak{r} of \mathfrak{B} is also a minimal right ideal of $K[\mathfrak{S}_m]$, and the \dagger idempotent element ε which generates \mathfrak{r} is a \dagger primitive idempotent of $K[\mathfrak{S}_m]$. From the theory of symmetric groups (\rightarrow 362 Representations H) the primitive idempotents of $K[\mathfrak{S}_m]$ are all given (up to isomorphism) by \dagger Young's diagrams $T(f_1, \dots, f_k)$ ($f_1 \geq f_2 \geq \dots \geq f_k > 0$, $m = f_1 + \dots + f_k$). In this setting, we have

(4) Let $\varepsilon = \varepsilon(f_1, \dots, f_k)$ be the primitive idempotent determined by Young's diagram $T(f_1, \dots, f_k)$. Then $\varepsilon K[\mathfrak{S}_m] \subset \mathfrak{B}$ if and only if $k \leq n$. In this case, put $\varepsilon K[\mathfrak{S}_m] = \mathfrak{r}$, $\mathfrak{r}(V^m) =$

$\varepsilon(V^m) = V^m(T(f_1, \dots, f_k))$ and denote the irreducible representation of $GL(V)$ on $V^m(T(f_1, \dots, f_k))$ by $A \rightarrow D(A; f_1, \dots, f_k)$. We call (f_1, \dots, f_k) the **signature** of this irreducible representation.

(5) The representation $D(A; f_1, \dots, f_k)$ is an irreducible polynomial representation of $GL(V)$. Furthermore, for any irreducible polynomial representation ρ of $GL(V)$, there exists a unique $D(A; f_1, \dots, f_k)$ equivalent to ρ . For example, if $k = 1$, then $f_1 = m$, $\varepsilon = (m!)^{-1} \sum_{\sigma \in \mathfrak{S}_m} \sigma$, and $V^m(T(m))$ is the space of \dagger symmetric tensors of degree m . If $f_1 = \dots = f_k = 1$, then $k = m$, $\varepsilon = (m!)^{-1} \sum_{\sigma \in \mathfrak{S}_m} (\text{sgn } \sigma) \sigma$, and $V^m(T(1, \dots, 1))$ is the space of \dagger alternating tensors of degree m .

(6) Let $\chi(A; f_1, \dots, f_k)$ be the \dagger character of the irreducible representation $D(A; f_1, \dots, f_k)$. Then

$$\chi(A; f_1, \dots, f_k) = \begin{vmatrix} \varepsilon_1^{l_1} & \varepsilon_1^{l_2} & \dots & \varepsilon_1^{l_n} \\ \dots & \dots & \dots & \dots \\ \varepsilon_n^{l_1} & \varepsilon_n^{l_2} & \dots & \varepsilon_n^{l_n} \end{vmatrix} \div \begin{vmatrix} \varepsilon_1^{n-1} & \varepsilon_1^{n-2} & \dots & \varepsilon_1 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \varepsilon_n^{n-1} & \varepsilon_n^{n-2} & \dots & \varepsilon_n & 1 \end{vmatrix},$$

where $\varepsilon_1, \dots, \varepsilon_n$ are the \dagger eigenvalues of A and $l_1 = f_1 + (n-1)$, $l_2 = f_2 + (n-2)$, \dots , $l_n = f_n$ (set $f_{k+1} = \dots = f_n = 0$). Hence the degree d of $D(A; f_1, \dots, f_k)$ is expressed as

$$d = D(l_1, \dots, l_n) / D(n-1, \dots, 1, 0),$$

where $D(x_1, \dots, x_n) = \prod_{i < j} (x_i - x_j)$.

(7) In particular, denote the character of $D(A; m)$ by $p_m = p_m(A)$. Then they satisfy $|I - zA|^{-1} = p_0 + p_1 z + p_2 z^2 + \dots$ and

$$\chi(A; f_1, \dots, f_k) = \begin{vmatrix} p_{f_1} & p_{f_1+1} & \dots & p_{f_1+(n-1)} \\ p_{f_2-1} & p_{f_2} & \dots & p_{f_2+(n-2)} \\ \dots & \dots & \dots & \dots \\ p_{f_n-(n-1)} & p_{f_n-(n-2)} & \dots & p_{f_n} \end{vmatrix},$$

where we put $f_{k+1} = f_{k+2} = \dots = f_n = 0$, $p_{-1} = p_{-2} = \dots = 0$. This matrix is simply written as $|p_{l_i - (n-1)}, \dots, p_{l_i}|$, with the convention that in each row, we set $l_1 = f_1 + (n-1)$, \dots , $l_{n-1} = f_{n-1} + 1$, $l_n = f_n$.

E. Determination of the Rational Representations of $SL(V)$

The rational representations of $SL(V)$ are completely reducible. By restricting any irreducible representation $D(A; f_1, \dots, f_n)$ ($f_1 \geq f_2 \geq \dots \geq f_n \geq 0$) of $GL(V)$ to $SL(V)$, we get an irreducible rational representation $\tilde{D}(A; f_1, \dots, f_n)$ of $SL(V)$. Furthermore, any irreducible rational representation of $SL(V)$

can be obtained in this way. $\tilde{D}(A; f_1, \dots, f_n)$ and $\tilde{D}(A; f'_1, \dots, f'_n)$ are equivalent representations of $SL(V)$ if and only if $f_i - f_{i+1} = f'_i - f'_{i+1}$ ($i = 1, \dots, n-1$).

F. Unitary Groups

The set $U(n)$ of all $n \times n$ \dagger unitary matrices with complex elements is a group under multiplication (\rightarrow 269 Matrices). This group $U(n)$ is called the **unitary group** (or **unitary transformation group**) of degree n . The subset of $U(n)$ consisting of all matrices of determinant 1 is a normal subgroup of $U(n)$. This group is called the **special unitary group** and is denoted by $SU(n)$.

$U(n)$ and $SU(n)$ are subgroups of $GL(n, \mathbb{C})$ and $SL(n, \mathbb{C})$, respectively, and can be obtained from these groups through the \dagger unitary restriction. Hence they are both compact, connected Lie groups; in particular, $SU(1)$ is composed only of the identity and $U(1)$ is the multiplicative group of all complex numbers of absolute value 1. The center \mathfrak{z} of $U(n)$ is the set of all diagonal matrices λI ($\lambda \in \mathbb{C}, |\lambda| = 1$), and we have

$$\mathfrak{z} \cong U(1), \quad \mathfrak{z} \cdot SU(n) = U(n),$$

$$U(n)/SU(n) \cong U(1).$$

Moreover, for $n \geq 2$, $SU(n)$ is a simple, semi-simple, and simply connected Lie group, which gives one of four infinite series of simple compact Lie groups.

$U(n)/\mathfrak{z}$ is denoted by $PU(n)$ and is called the **projective unitary group**. We have the relations $PU(n) \cong SU(n)/\mathfrak{z} \cap SU(n)$, $\mathfrak{z} \cap SU(n) \cong \mathbb{Z}/n\mathbb{Z}$. Hence $PU(n)$ is locally isomorphic to $SU(n)$.

G. Irreducible Representations of $U(n)$

Restricting the irreducible representation $D(A; f_1, \dots, f_k)$ of $GL(n, \mathbb{C})$ on $SU(n)$, we obtain a continuous irreducible representation of $SU(n)$, and conversely, all continuous irreducible representations of $SU(n)$ are obtained in this manner. Similarly, any continuous irreducible representations of $U(n)$ are given by $A \rightarrow |A|^e D(A; f_1, \dots, f_k)$, where e is an integer. Since both $U(n)$ and $SU(n)$ are compact, any continuous representation of these groups can be decomposed into a direct sum of the irreducible representations mentioned above (\rightarrow 69 Compact Groups).

The representation theory of $U(n)$ and $SU(n)$ is important as the most typical and concrete example of the representation theory of general compact Lie groups (\rightarrow 69 Compact Groups, 248 Lie Algebras, 249 Lie Groups).

H. Unitary Groups over General Fields

A unitary matrix and the unitary group can also be defined over some fields other than the field \mathbb{C} of complex numbers. Namely, let P be a field and K a quadratic extension field of P ; for an element ξ of K , let $\bar{\xi}$ be the \dagger conjugate of ξ over P . Then a matrix of degree n with entries in K is called a unitary matrix of K (relative to P) if it leaves invariant the \dagger Hermitian form $\xi_1 \bar{\xi}_1 + \xi_2 \bar{\xi}_2 + \dots + \xi_n \bar{\xi}_n$. The multiplicative group consisting of all unitary matrices is called the **unitary group over K** (relative to P) and is denoted by $U(n, K, P)$; its subgroup consisting of all unitary matrices of determinant 1 is called the **special unitary group over K** and is denoted by $SU(n, K, P)$. The quotient group of $SU(n, K, P)$ by its subgroup consisting of all λI ($\lambda^n = 1, |\lambda| = 1$) is called the **projective special unitary group over K** and is denoted by $PSU(n, K, P)$. In particular, when K and P are the finite fields \mathbb{F}_{q^2} and \mathbb{F}_q ($q = p^m$), $U(n, K, P)$, $SU(n, K, P)$, $PSU(n, K, P)$ are written simply as $U(n, q)$, $SU(n, q)$, $PSU(n, q)$. Then for $n \geq 3$, each $PSU(n, q)$ is a noncommutative simple group, except for $PSU(3, 2)$ (\rightarrow 151 Finite Groups I).

I. Orthogonal Groups

The set of all \dagger orthogonal matrices of degree n (with real entries) forms a group under multiplication. This group $O(n)$ is called the **orthogonal group** (or **orthogonal transformation group**) of degree n . The subset of $O(n)$ consisting of all orthogonal matrices of determinant 1 forms a normal subgroup of $O(n)$ of index 2. This group $SO(n)$ (also denoted by O_n^+) is called the **rotation group** (**special orthogonal group** or **proper orthogonal group**) of degree n . Geometrically, $O(n)$ is the set of all orthogonal transformations leaving a point in Euclidean space of dimension n fixed, and $SO(n)$ is composed of all rotations around the point.

Both $O(n)$ and $SO(n)$ are compact Lie groups, and $SO(n)$ coincides with the connected component of $O(n)$ which contains the identity. For $n = 3$ or $n \geq 5$, each $SO(n)$ is a simple and semisimple Lie group. Following the theory of Lie algebras, we divide the set of all $SO(n)$ ($n \geq 3$ but $n \neq 4$) into two classes according as n is even or odd, and we thus get two of the four infinite series of simple and semisimple compact Lie groups (for $SO(4)$, for example, see [1]).

Although $SO(n)$ ($n \geq 3$) is a connected Lie group, it is not simply connected. The simply connected Lie group which is locally isomorphic to $SO(n)$ is called the **spinor group** and is denoted by $Spin(n)$. $SO(n)$ is isomorphic to the

quotient group of $Spin(n)$ by a normal subgroup of order 2. Let z be the center of $Spin(n)$. Then $z \cong \mathbb{Z}/2\mathbb{Z}$ for odd n , $z \cong \mathbb{Z}/4\mathbb{Z}$ for $n \equiv 2 \pmod{4}$, and $z \cong (\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/2\mathbb{Z})$ for $n \equiv 0 \pmod{4}$ (\rightarrow 61 Clifford Algebras).

The group $O(n, \mathbb{C})$ of all \dagger complex orthogonal matrices is called the **complex orthogonal group**, and the group $SO(n, \mathbb{C})$ of all matrices in $O(n, \mathbb{C})$ of determinant 1 is called the **complex special orthogonal group**. $SO(n, \mathbb{C})$ ($n \geq 3, n \neq 4$) is a simple and semisimple complex Lie group.

J. Irreducible Representations of Orthogonal Groups

In the same way as for $GL(n, K)$, the irreducible representations of $O(n)$ can be obtained by decomposing the tensor product $D_m(A) = A \otimes \dots \otimes A$ of m copies of an orthogonal matrix A using Young's diagram. Namely, consider the Young's diagram $T(f_1, f_2, \dots, f_k)$ such that the sum of the lengths of the first column and of the second column is not greater than n , and call it an $O(n)$ diagram. Then to any $O(n)$ diagram $T = T(f_1, \dots, f_k)$, there corresponds an absolutely irreducible representation $D^0(A; f_1, f_2, \dots, f_k)$, and the representations $D^0(A; f_1, f_2, \dots, f_k)$ corresponding to two distinct $O(n)$ diagrams are mutually inequivalent. $D_m(A)$ can be decomposed into the direct sum of those representations $D^0(A; f_1, f_2, \dots, f_k)$ such that $f = f_1 + \dots + f_k$ takes the values $m, m-2, m-4, \dots$. Furthermore, any continuous irreducible representation of $O(n)$ is equivalent to a $D^0(A; f_1, f_2, \dots, f_k)$ obtained from some $O(n)$ diagram $T = T(f_1, f_2, \dots, f_k)$.

In general, two $O(n)$ diagrams T and T' are called **mutually associated diagrams** if the sum of the lengths of their first columns is equal to n and if the lengths of each column other than the first one coincide. In particular, if $T = T(f_1, f_2, \dots, f_k)$ and $2k = n$, then T is said to be self-associated. The set of all $O(n)$ diagrams can be divided into pairs of mutually associated T, T' (and self-associated $T = T'$). Suppose that we are given mutually associated diagrams T and T' and that the length k of the first column of $T = T(f_1, f_2, \dots, f_k)$ is not greater than $n/2$. Then the character $\chi_T(A)$ of $D^0(A; f_1, f_2, \dots, f_k)$ corresponding to T and the character $\chi_{T'}(A)$ of the irreducible representation corresponding to T' are given by

$$\chi_T(A) = |p_{l-(v-1)} - p_{l-(v+1)}|,$$

$$p_{l-(v-2)} - p_{l-(v+2)}, \dots, p_l - p_{l-2v},$$

$$\chi_{T'}(A) = |A| \chi_T(A), \quad v = [n/2],$$

where p_i and $|p_{l-(v-1)} - p_{l-(v+1)}, \dots|$ have the same meaning as in the formula for the char-

acters of irreducible representations of $GL(n, K)$.

The irreducible representations of $SO(n)$ can be obtained immediately from those of $O(n)$. Namely, if $T = T(f_1, f_2, \dots, f_k)$ is not self-associated, $D^0(A; f_1, f_2, \dots, f_k)$ is irreducible as a representation of $SO(n)$, and the representations of $SO(n)$ derived from T and the associated T' coincide. If T is self-associated, $D^0(A; f_1, f_2, \dots, f_k)$ can be decomposed into two irreducible representations of $SO(n)$ of the same degree over the field of complex numbers. Furthermore, the irreducible representations of $SO(n)$ obtained in this way from different pairs of associated diagrams are mutually inequivalent, while any continuous irreducible representation of $SO(n)$ is equivalent to one of these representations. For the representations of $SO(3)$ (the rotation group of degree 3) \rightarrow 353 Racah Algebra.

Since $SO(n)$ is isomorphic to the quotient group of $Spin(n)$ by a normal subgroup N of order 2, a continuous representation of $Spin(n)$ which is not the identity representation on N can be considered as a double-valued representation of $SO(n)$. This representation is called the **spin representation** and is important in the field of applied mathematics.

The orthogonal group $O(n)$ consists of all $n \times n$ real matrices which leave invariant the quadratic form $\xi_1^2 + \dots + \xi_n^2$, while the group of all $n \times n$ real matrices which leave invariant the quadratic form $\xi_1^2 + \dots + \xi_r^2 - \xi_{r+1}^2 - \dots - \xi_n^2$ of \dagger signature $(r, n-r)$ is called the **Lorentz group** of signature $(r, n-r)$. The case for $n=4$ and $r=3$ is used in special relativity (\rightarrow 359 Relativity). Let G_0 be the connected component of the identity of the Lorentz group of signature $(3, 1)$. Then G_0 is called the **proper Lorentz group**. For $\sigma = (g_{ij}) \in G$, we have $|\sigma| = \pm 1$ and $g_{44} \geq 1$ or $g_{44} \leq -1$. Moreover, we have $G_0 = \{\sigma \mid |\sigma| = 1, g_{44} \geq 1\}$, $G/G_0 \cong (\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/2\mathbb{Z})$ (\dagger four group), and $G_0 \cong SL(2, \mathbb{C})/\{\pm I\}$.

K. Orthogonal Groups over General Fields

Orthogonal groups can also be defined over other general fields than the field of real numbers as follows: Fix a \dagger quadratic form $Q(\xi, \xi) = \sum_{i,j=1}^n \alpha_{ij} \xi_i \xi_j$ ($|\alpha_{ij}| \neq 0$) over a field K . Then a linear transformation of ξ_i ($i=1, 2, \dots, n$) over K which leaves Q invariant is called an **orthogonal transformation** with respect to Q . The set of all orthogonal transformations forms a group. This group is denoted by $O(n, K, Q)$ or simply $O(Q)$ and is called the **orthogonal (transformation) group over K with respect to Q** . In particular, the normal subgroup of all transformations in $O(n, K, Q)$ of determinant 1 is called the **special orthogonal group over K with**

respect to Q and is denoted by $SO(n, K, Q)$ (or simply $SO(Q)$). $O(n)$ and $SO(n)$ are special cases of $O(n, K, Q)$ and $SO(n, K, Q)$, where K is the field \mathbf{R} of real numbers and $Q(\xi, \xi)$ is the unit quadratic form $\xi_1^2 + \xi_2^2 + \dots + \xi_n^2$.

Let $\Omega(n, K, Q)$ be the \dagger commutator subgroup of $O(n, K, Q)$. Then this subgroup coincides with the commutator subgroup of $SO(n, K, Q)$. If K is of characteristic $\neq 2$, and if $n \geq 5$ and the \dagger index v of $Q \geq 1$, then $\Omega(n, K, Q)/3$ (3 is the center of $\Omega(n, K, Q)$) is a simple group, where $3 = \{I\}$ or $3 = \{\pm I\}$ (L. Dickson, J. Dieudonné). Suppose that K is a finite field F_q (of characteristic $\neq 2$). Then we have $v = m$ if $n = 2m + 1$, and $v = m$ or $m - 1$ if $n = 2m$. Hence $v \geq 2$ if $n \geq 5$. If $v = 0$ and $K = \mathbf{R}$, we have $\Omega(n, \mathbf{R}, Q) = SO(n)$ and, as mentioned before, $SO(n)/3$ is simple for $n \geq 5$. The same proposition also holds when K is an \dagger algebraic number field (M. Kneser, 1956). If K is of characteristic 2, then $O(n, K, Q) = SO(n, K, Q)$, $3 = \{I\}$, and $\Omega(n, K, Q)$ is a simple group in many cases (Dieudonné [5]). For the case where K is a finite field (Dickson) \rightarrow 151 Finite Groups I.

L. Symplectic Groups

Let $\xi_1, \xi_2, \dots, \xi_{2n}$ and $\eta_1, \eta_2, \dots, \eta_{2n}$ be two sets of variables, and suppose that the same linear transformation A over a field K acts on them (from the left). If A leaves the \dagger bilinear form $\sum_{i=1}^n (\xi_{2i-1}\eta_{2i} - \xi_{2i}\eta_{2i-1})$ invariant, this linear transformation (or the corresponding matrix) A is called a **symplectic transformation (symplectic matrix)** of degree $2n$. The set of all symplectic transformations (or matrices) of degree $2n$ over K forms a group denoted by $Sp(n, K)$ and called the **symplectic group (symplectic transformation group, complex group, or Abelian linear group) over K** .

Any matrix in $Sp(n, K)$ is always of determinant 1, and the center 3 of $Sp(n, K)$ consists of I and $-I$. The quotient group $PSp(n, K)$ of $Sp(n, K)$ by 3 is called the **projective symplectic group over K** . Except for the three cases $n = 1$, $K = F_2$; $n = 1$, $K = F_3$; and $n = 2$, $K = F_2$, the group $PSp(n, K)$ ($n \geq 1$) is always simple.

Properties of Symplectic Groups as Lie Groups.

When K is the field \mathbf{C} of complex numbers or the field \mathbf{R} of real numbers, $Sp(n, K)$ is a Lie group. The intersection of the **complex symplectic group** $Sp(n, \mathbf{C})$ and the unitary group $U(2n)$, namely, the unitary restriction of $Sp(n, \mathbf{C})$, is denoted by $Sp(n)$ and is called the **unitary symplectic group** (or simply **symplectic group**). $Sp(n, \mathbf{C})$ is a simple and semisimple complex Lie group, and both $Sp(n, \mathbf{R})$ and $Sp(n)$ are simple and semisimple Lie groups. Moreover, $Sp(n)$ is compact and simply con-

nected and gives one of the four series of simple, semisimple and compact Lie groups (\rightarrow 249 Lie Groups).

Let \mathbf{H}^n be the linear space of dimension n over the \dagger quaternion field \mathbf{H} . Define the inner product of two elements $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in \mathbf{H}^n by $(x, y) = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n$ (\bar{y}_i is the \dagger conjugate quaternion of y_i), and consider the group of all linear transformations which leave this inner product invariant. Then this group is isomorphic to $Sp(n)$. $Sp(n)$ is thus compared with the orthogonal group $O(n)$, which leaves invariant the inner product of a linear space over the field \mathbf{R} of real numbers and with the unitary group $U(n)$, which has the same property over the field \mathbf{C} of complex numbers (C. Chevalley [4, ch. 1]).

M. Irreducible Representations of Symplectic Groups

In the same way as for $GL(n, K)$, the representation $D_m(A) = A \otimes \dots \otimes A$ (tensor product of m copies of A) of $Sp(n, \mathbf{C})$ can be decomposed into irreducible components using Young's diagram. Namely, for any Young's diagram $T = T(f_1, f_2, \dots, f_k)$ ($k \leq n$) such that the number k of rows is not greater than n , an irreducible representation $D^s(A; f_1, \dots, f_k)$ of $Sp(n, \mathbf{C})$ is determined. These $D^s(A; f_1, \dots, f_k)$ are mutually inequivalent, and $D_m(A)$ can be decomposed into the direct sum of representations $D^s(A; f_1, f_2, \dots, f_k)$ such that $f = f_1 + \dots + f_k$ is equal to any of the values $m, m-2, m-4, \dots$. The character of $D^s(A; f_1, f_2, \dots, f_k)$ is given by

$$\chi_T(A) = |p_{l-n+1}, p_{l-n+2}, p_{l-n}, \dots, p_l + p_{l-2n+2}|,$$

where p_i and $|p_{l-n+1}, p_{l-n+2}, p_{l-n}, \dots|$ have the same meaning as in the formula for the characters of the irreducible representations of $GL(n, K)$.

For the matrices A in $Sp(n)$, $D^s(A; f_1, f_2, \dots, f_k)$ gives rise to a continuous irreducible representation of $Sp(n)$. Furthermore, any continuous irreducible representation of $Sp(n)$ is equivalent to a representation $D^s(A; f_1, f_2, \dots, f_k)$ corresponding to some diagram T .

N. Relations among Various Classical Groups

There are some isomorphisms (homomorphisms) among the classical groups mentioned above. For general fields $K \rightarrow$ [1, 5]. For finite fields $K \rightarrow$ 151 Finite Groups I. For $K = \mathbf{R}$ or \mathbf{C} , the following isomorphisms hold: $SO(3) \cong SU(2)/\{\pm I\}$, $SU(2) \cong Sp(1)$, $SO(5) \cong Sp(2)/\{\pm I\}$, $SO(6) \cong SO(4)/\{\pm I\}$ (\rightarrow 248 Lie Algebras, 249 Lie Groups).

O. Classical Groups over Noncommutative Fields

Let V be a right linear space over a noncommutative field K . Then the set of all linear transformations of V forms a group under the multiplication defined by the composition of mappings. This group $GL(V)$ is called the **general linear group** on V . It is isomorphic to the multiplicative group of all $n \times n$ invertible matrices with entries in K . The commutator subgroups $SL(V)$ and $SL(n, K)$ of $GL(V)$ and $GL(n, K)$, respectively, are called the **special linear group** of degree n on V and over K , respectively. Now, suppose that an element A of $GL(V)$ leaves each element of a subspace U of dimension $n-1$ of V fixed. Choose an element x of V which does not belong to U , and set $Ax \equiv x\alpha \pmod{U}$; $\alpha \in K$ depends not only on A but also on the choice of x . However, the conjugate class $\dot{\alpha} = \{\lambda\alpha\lambda^{-1} \mid \lambda \in K^*\}$ of α in the multiplicative group K^* of K is determined only by A . In particular, if $\dot{\alpha} = \{1\}$ and $A \neq I$, then A is called a **transvection**. For a \dagger matrix unit E_{ij} , $B_{ij}(\alpha) = I + \alpha E_{ij}$ is a transvection if $i \neq j$ and $\alpha \neq 0$. $SL(V)$ coincides with the subgroup of $GL(V)$ generated by all transvections. This fact also holds when K is a commutative field, except when $n=2$ and $K = \mathbb{F}_2$. In this case, transvections generate the whole $GL(2, 2)$, which is isomorphic to the symmetric group \mathfrak{S}_3 of degree 3 and does not coincide with the commutator subgroup. The center \mathfrak{z} of $GL(n, K)$ consists of all scalar matrices corresponding to nonzero elements in the center of K . Let C be the commutator subgroup of the multiplicative group K^* of K . Then for $n \geq 2$, $GL(n, K)/SL(n, K)$ is isomorphic to K^*/C . This isomorphism can be obtained by appropriately defining, for $A \in GL(n, K)$, an element $\det A$ of K^*/C which is called the **determinant** of A [6, 9]. The center \mathfrak{z}_0 of $SL(n, K)$ is $\{\alpha I \mid \alpha^n \in C\}$. The quotient group $PSL(n, K) = SL(n, K)/\mathfrak{z}_0$ is called the **projective special linear group** of degree n over K . If K is a noncommutative field, then $PSL(n, K)$ ($n \geq 2$) is always a simple group [5, 8].

Next, let K be any field (commutative or noncommutative), and let V be a right linear space of dimension n over K . Consider a Hermitian form $f(x, y)$ (\rightarrow 256 Linear Spaces) on V relative to an \dagger involution J of K . If for a fixed element ε in the center of K we have $f(x, y) = \varepsilon f(y, x)$, then f is called an **ε -Hermitian form**. For the rest of this article, f is assumed to be an ε -Hermitian form on V . Let W be a subspace of V . If $f(x, y) = 0$ for any $x, y \in W$, then W is called a **totally isotropic subspace**. The largest dimension m of the totally isotropic subspaces of V is called the **index** of f . We always have $2m \leq n$. If $f(Ax, Ay) = f(x, y)$

for any $x, y \in V$, then A is called a **unitary transformation** relative to f .

The set $U(n, K, f)$ of all unitary transformations relative to f forms a subgroup of $GL(V)$. This group is called the **unitary group** relative to f . Also, the group $SU(n, K, f) = U(n, K, f) \cap SL(n, K)$ is called the **special unitary group**. When $J = 1$ and $\varepsilon = 1$, a unitary transformation (unitary group) is called an **orthogonal transformation (orthogonal group)**, and $U(n, K, f)$ is written as $O(n, K, f)$. Also, when $J = 1$ and $\varepsilon = -1$, a unitary transformation (unitary group) is called a **symplectic transformation (symplectic group)**, and $U(n, K, f)$ is written as $Sp(n, K)$. In fact, in these cases, for arbitrary choice of f , the corresponding groups are mutually isomorphic.

An ε -Hermitian form f is called an **ε -trace form** if for any $x \in V$, there exists an $\alpha \in K$ which satisfies $f(x, x) = \alpha + \varepsilon\alpha^J$. If $J = 1$, $\varepsilon = -1$ (hence K is commutative) or $\varepsilon = 1$ and K is of characteristic $\neq 2$, then any ε -Hermitian form is an ε -trace form. If f is an ε -trace form, a linear mapping B of any subspace W of V into V such that for any $x, y \in W$, $f(Bx, By) = f(x, y)$ can be extended to an element A of the unitary group $U(n, K, f)$ relative to f (**Witt's theorem**). In particular, $U(n, K, f)$ acts transitively on the maximal totally isotropic subspaces, and their dimensions are equal to the index m of f . Now, let P be a **Pythagorean ordered field** (an ordered field which contains square roots of any positive element). If $K = P$ and $J = 1$, or if $K = P(\sqrt{-1})$, or if the noncommutative field K is a \dagger quaternion algebra over P and J is the operation of \dagger conjugation of K , then for two Hermitian forms f, f' , their unitary groups $U(n, K, f)$ and $U(n', K, f')$ are isomorphic if and only if $n = n'$, and the indices of f and f' are equal. In this case, $U(n, K, f)$ can be written as $U(n, m, K)$, where m is the index of f . If the field K is a \dagger quaternion algebra over P and f is an \dagger anti-Hermitian form, there exists an orthogonal basis (e_i) of V such that $f(e_i, e_i) = j$ (quaternion unit), $1 \leq i \leq n$. Hence, in this case, the unitary group $U(n, K, f)$ relative to f is determined only by n and K .

Suppose that we are given an ε -trace form f over a general field K whose index m is not equal to 0. We exclude the case where $J = 1$ and $\varepsilon = 1$. Then the unitary group $U(n, K, f)$ contains transvections. Let $T(n, K, f)$ denote the subgroup of $U(n, K, f)$ generated by transvections which are unitary transformations. If $m \geq 2$, then $T(n, K, f)$ is the commutator subgroup of $U(n, K, f)$. The center W_n of $T(n, K, f)$ coincides with the intersection of $T(n, K, f)$, and the center \mathfrak{z} of $GL(n, K)$. If $n \geq 3$ and K contains more than 25 elements, then the quotient group $T(n, K, f)/W_n$ is a simple group [6]. Also, if K is commutative and $n \geq 2$, $m \geq 1$,

$J \neq 1$, then $T(n, K, f) = SU(n, K, f)$, except for the case where $n = 3$, $K = \mathbf{F}_4$.

If K is the field \mathbf{R} of real numbers, the field \mathbf{C} of complex numbers, or the quaternion field \mathbf{H} , then $GL(n, K)$, $SL(n, K)$, and $U(n, K, f)$ are all Lie groups. In particular, $SL(n, K)$ and $U(n, K, f)$ are simple Lie groups except in the following three cases: (1) $n = 1$, $K = \mathbf{R}$ or \mathbf{C} ; (2) $n = 2$, $K = \mathbf{R}$, $J = 1$, $\varepsilon = 1$; (3) $n = 4$, $K = \mathbf{R}$ or \mathbf{C} , $J = 1$, $\varepsilon = 1$, $m = 2$. In cases (1) and (2) they are commutative groups, and in case (3) they are locally direct sums of two noncommutative simple groups.

Suppose that $K = \mathbf{H}$. Since \mathbf{H} contains \mathbf{C} as a subfield, a vector space V of dimension n over \mathbf{H} has the structure of a vector space of dimension $2n$ over \mathbf{C} . From this fact, $GL(n, \mathbf{H})$ can be considered as a subgroup of $GL(2n, \mathbf{C})$ in a natural way.

Each of the complex classical simple groups $G = SL(n, \mathbf{C})$, $SO(n, \mathbf{C})$, $Sp(n, \mathbf{C})$ has the structure of an \dagger algebraic group defined over \mathbf{R} (\rightarrow 13 Algebraic Groups). The **real forms** of G , i.e., the algebraic subgroups of G whose scalar extension to \mathbf{C} is G , can be realized as $SL(n, K)$, $U(n, K, f)$ corresponding to $K = \mathbf{R}, \mathbf{C}, \mathbf{H}$. Namely, a real form of a complex classical group G is conjugate in G to one of the following groups: (i) The real forms of $SL(n, \mathbf{C})$: $SL(n, \mathbf{R})$ (type AI); $SL(k, \mathbf{H})$ only for $n = 2k$ (type AII); and the special unitary group $SU(n, m, \mathbf{C})$, $0 \leq m \leq [n/2]$, relative to a Hermitian form f of index m (type AIII). (ii) The real forms of $SO(2n+1, \mathbf{C})$: the proper orthogonal group $SO(2n+1, m, \mathbf{R})$, $0 \leq m \leq n$, relative to a quadratic form of index m on a space of dimension $2n+1$ (type BI and BII). (iii) The real forms of $SO(2n, \mathbf{C})$: $SO(2n, m, \mathbf{R})$, $0 \leq m \leq n$ (type DI and DII); and $U(n, H, f)$ relative to an anti-Hermitian form f on \mathbf{H} (type DIII). (iv) The real forms of $Sp(n, \mathbf{C})$: $Sp(n, \mathbf{R})$ (type CI); the unitary group $U(2n, m, \mathbf{H})$, $0 \leq m \leq n$, relative to a Hermitian form f of index m on \mathbf{H} (type CII); and $Sp(n)$ corresponds to the special case $m = 0$. The quotient groups of these real forms by their centers can all be realized as the groups of automorphisms of semisimple algebras with involutions J which commute with J (A. Weil [10]).

References

- [1] B. L. van der Waerden, Gruppen von linearen Transformationen, Erg. Math., Springer, 1935 (Chelsea, 1948).
- [2] H. Weyl, The classical groups, Princeton Univ. Press, 1939, revised edition, 1946.
- [3] N. Bourbaki, Eléments de mathématique, Algèbre, ch. 9, Actualités Sci. Ind., 1272a, Hermann, 1959.

- [4] C. Chevalley, Theory of Lie groups I, Princeton Univ. Press, 1946.
- [5] J. Dieudonné, Sur les groupes classiques, Actualités Sci. Ind., Hermann, 1948.
- [6] J. Dieudonné, La géométrie des groupes classiques, Springer, 1955.
- [7] C. Chevalley, The algebraic theory of spinors, Columbia Univ. Press, 1954.
- [8] M. Eichler, Quadratische Formen und orthogonale Gruppen, Springer, 1952.
- [9] E. Artin, Geometric algebra, Interscience, 1957.
- [10] A. Weil, Algebras with involutions and the classical groups, J. Indian Math. Soc., 24 (1960), 589–623.

61 (III.16) Clifford Algebras

A. Definitions and Basic Properties

Let V be an n -dimensional \dagger linear space over a field K , and let Q be a \dagger quadratic form on V . Denote the \dagger tensor algebra over V by $T(V)$, the tensor multiplication by \otimes . Let $I(Q)$ be the two-sided ideal of $T(V)$ generated by the elements $x \otimes x - Q(x) \cdot 1$ ($x \in V$). The resulting \dagger quotient associative algebra $T(V)/I(Q)$ is then denoted by $C(Q)$ and is called the **Clifford algebra** of the quadratic form Q . The elements of $C(Q)$ are called **Clifford numbers**.

The composite of two canonical mappings $\tau: V \rightarrow T(V)$, $\sigma: T(V) \rightarrow C(Q)$ is a linear injection $\sigma \circ \tau: V \rightarrow C(Q)$. Hence we can regard V as a linear subspace of $C(Q)$ via $\sigma \circ \tau$. Then $C(Q)$ is an associative algebra over K generated by 1 and V . Furthermore, $x^2 = Q(x) \cdot 1$ for every x in V .

Indeed, $C(Q)$ is the universal associative algebra with these properties. That is, let A be any associative algebra with a unity element, and let $f: V \rightarrow A$ be a linear mapping such that $f(x)^2 = Q(x) \cdot 1$ for every x in V . Then f can be extended uniquely to an algebra homomorphism $\tilde{f}: C(Q) \rightarrow A$ with $\tilde{f}(1) = 1$. Furthermore, let Φ be the \dagger symmetric bilinear form associated with Q : $\Phi(x, y) = Q(x+y) - Q(x) - Q(y)$, $x, y \in V$. Then $xy + yx = \Phi(x, y) \cdot 1$ for every x, y in V . $C(Q)$ is of dimension 2^n over K . If e_1, \dots, e_n is a basis of V , then

$$1, e_i, e_i e_j (i < j), \dots, e_1 e_2 \dots e_n$$

form a basis of $C(Q)$. In particular, if $\{e_i\}$ is an orthogonal basis relative to Q , we have

$$e_i e_j = -e_j e_i, \quad e_i^2 = Q(e_i) \cdot 1;$$

$$i, j = 1, \dots, n, \quad i \neq j. \quad (1)$$

In this case, $C(Q)$ can be defined as an associative algebra (with a unity element) generated by the $\{e_i\}$ together with the defining relations (1). In particular, for $Q=0$, $C(Q)$ is the exterior algebra (\dagger Grassmann algebra) over V .

B. The Principal Automorphism and the Principal Antiautomorphism of $C(Q)$

There exists a unique automorphism α of the algebra $C(Q)$ such that $\alpha(x) = -x$ for every x in V . This automorphism α is called the **principal automorphism** of $C(Q)$, and we have $\alpha^2 = 1$.

Also, there exists a unique antiautomorphism β of the algebra $C(Q)$ such that $\beta(x) = x$ for every x in V . This antiautomorphism β is called the **principal antiautomorphism** of $C(Q)$, and we have $\beta^2 = 1$.

For the rest of this article we assume that the \dagger discriminant of Q is $\neq 0$. We also assume for the sake of simplicity that the characteristic of K is $\neq 2$. Let $C^+ = C^+(Q) = K \cdot 1 + V^2 + V^4 + \dots$, and $C^- = C^-(Q) = V + V^3 + V^5 + \dots$. Then $C(Q)$ is the direct sum of the linear subspaces $C^+(Q)$ and $C^-(Q)$. Furthermore, $C^+C^+ \subset C^+$, $C^+C^- \subset C^-$, $C^-C^+ \subset C^-$, and $C^-C^- \subset C^+$. Thus $C(Q) = C^+ + C^-$ has the structure of a \dagger graded algebra with the index group $\{\pm 1\}$, and C^+ is a subalgebra of $C(Q)$. The elements of $C^+(Q)$, $C^-(Q)$ are called **even elements** and **odd elements**, respectively. We have $\dim C^+(Q) = \dim C^-(Q) = 2^{n-1}$.

C. The Structure of $C(Q)$ and $C^+(Q)$

$C(Q)$ and $C^+(Q)$ are both \dagger separable, \dagger semi-simple associative algebras over K . Suppose n is even: $n = 2r$. Then $C(Q)$ is a \dagger simple algebra with K as its center; the center Z of $C^+(Q)$ is 2-dimensional over K . Let e_1, \dots, e_n be an orthogonal basis of V . Then 1 and $z = 2^r e_1 \dots e_n$ form a basis of Z , and we have

$$z^2 = 2^{2r}(-1)^r Q(e_1) \dots Q(e_n) = (-1)^r D,$$

where D is the \dagger discriminant of Φ relative to the basis $\{e_i\}$. Thus if $(-1)^r D$ has a square root in K , $Z \cong K \oplus K$ (direct sum), and so $C^+(Q)$ is decomposed into the direct sum of two simple algebras. If $(-1)^r D$ does not have a square root in K , then Z is a field and $C^+(Q)$ is a simple algebra. In particular, if the \dagger index of Q (i.e., the dimension of a maximal \dagger totally singular subspace of V (\rightarrow 348 Quadratic Forms)) is r , $C(Q)$ is isomorphic to the \dagger total matrix algebra of degree 2^r over K , and $C^+(Q)$ is isomorphic to the direct sum of two copies of the total matrix algebra of degree 2^{r-1} over K .

Now suppose that n is odd: $n = 2r + 1$. Then

$C^+(Q)$ is a simple algebra with K as its center. (In particular, if Q is of index r , then $C^+(Q)$ is isomorphic to the total matrix algebra of degree 2^r over K .) The center Z of $C(Q)$ is 2-dimensional over K , and we have $C(Q) \cong Z \otimes_K C^+(Q)$. If e_1, \dots, e_n is an orthogonal basis of V , then 1 and $z = e_1 \dots e_n$ form a basis of Z . Putting $z' = 2^{r+1} z$, we have $z'^2 = 2(-1)^r D$, where D is the discriminant of Φ relative to $\{e_i\}$. Thus if $2(-1)^r D$ has a square root in K , $C(Q)$ is the direct sum of two 2^{2r} -dimensional simple algebras. If $2(-1)^r D$ has no square root in K , then $C(Q)$ is a simple algebra.

D. The Clifford Group

Let G be the set of all invertible elements s in $C(Q)$ such that $sVs^{-1} = V$. Then G forms a group relative to the multiplication of $C(Q)$. This group G is called the **Clifford group** of the quadratic form Q . The subgroup $G^+ = G \cap C^+(Q)$ is called the **special Clifford group**. The linear transformation $\varphi(s): x \rightarrow sxs^{-1}$ of V induced by $s \in G$ belongs to the \dagger orthogonal group $O(Q)$ of V relative to Q . Moreover, the mapping $s \rightarrow \varphi(s)$ is a homomorphism from G into $O(Q)$. Thus φ is a \dagger representation of G on V . This representation φ is called the **vector representation** of G . The \dagger kernel of φ consists of invertible elements in the center Z of $C(Q)$. If $x \in G \cap V$, then $Q(x) \neq 0$ and $-\varphi(x)$ is the reflection mapping of V relative to the hyperplane orthogonal to x . If $n = \dim V$ is odd, $\varphi(G) = \varphi(G^+) = SO(Q)$. If n is even, $\varphi(G) = O(Q)$, $\varphi(G^+) = SO(Q)$.

Exploiting the principal antiautomorphism β of $C(Q)$, we obtain a homomorphism $N: G^+ \rightarrow K^*$ (the multiplicative group of K) defined by $N(s) = \beta(s)s$ ($s \in G^+$), and $N(s)$ is called the **spinorial norm** of $s \in G^+$. The normal subgroup of G^+ defined as the kernel of N is denoted by G_0^+ and is called the **reduced Clifford group** (of Q). The subgroup $\varphi(G_0^+)$ of $SO(Q)$ is denoted by $O_0^+(Q)$ and is called the **reduced orthogonal group**.

In particular, when the ground field K is the real number field \mathbf{R} , $O_0^+(Q)$ coincides with the \dagger identity component of the \dagger Lorentz group $O(Q)$. Furthermore, if Q is definite, $O_0^+(Q) \cong SO(n)$, so that the identity component $Spin(n)$ of G_0^+ is a \dagger simply connected \dagger covering group of $SO(n)$ via the covering homomorphism φ (with each point in $SO(n)$ covered twice). The group $Spin(n)$ is called the **spinor group** of degree n .

E. Spin Representations

In this section we assume that the ground field K is the complex number field \mathbf{C} and that n

$= \dim V \geq 3$. Then we have $O_0^+(Q) \cong SO(n, \mathbb{C})$, so G_0^+ is a simply connected covering group of $SO(n, \mathbb{C})$ via the covering homomorphism φ . In this section we denote G_0^+ by $Spin(n, \mathbb{C})$ and call it the **complex spinor group** of degree n . $Spin(n, \mathbb{C})$ is the \dagger complexification (\rightarrow 249 Lie Groups) of the compact Lie group $Spin(n)$ and is a complex analytic subgroup of the \dagger complex Lie group $C(Q)^*$ consisting of all invertible elements of $C(Q)$. With the bracket operation $[x, y] = xy - yx$, $C(Q)$ becomes the \dagger Lie algebra of $C(Q)^*$. Furthermore, the Lie subalgebra of $C(Q)$ associated with the complex analytic subgroup $Spin(n, \mathbb{C})$ is given by $\sum_{i < j} \mathbb{C} e_i e_j$, where e_1, \dots, e_n is an orthogonal basis of V . The spin representations of the group $Spin(n, \mathbb{C})$ are defined as follows:

(1) When n is odd: $n = 2r + 1$. Since $C^+(Q)$ is isomorphic to a total matrix algebra of degree 2^r over \mathbb{C} , $C^+(Q)$ has a unique (up to equivalence) \dagger irreducible representation $\tilde{\rho}$, which is of degree 2^r . The restriction of $\tilde{\rho}$ on $Spin(n, \mathbb{C})$ (on $Spin(n)$) defines an irreducible representation ρ of degree 2^r of $Spin(n, \mathbb{C})$ (of $Spin(n)$); ρ is called the **spin representation** of the group $Spin(n, \mathbb{C})$ (of $Spin(n)$). The elements in the representation space of ρ are called **spinors**. Thus we can say that a spinor is a quantity with 2^r components which obey the transformation law according to the spin representation (\rightarrow 258 Lorentz Group). This representation ρ defines a representation of the Lie algebra $\mathfrak{so}(n, \mathbb{C})$ of $Spin(n, \mathbb{C})$ (note that $\mathfrak{so}(n, \mathbb{C})$ is a \dagger complex simple Lie algebra of type B_r). This representation of $\mathfrak{so}(n, \mathbb{C})$ is also called the spin representation of $\mathfrak{so}(n, \mathbb{C})$. Note that ρ is not well defined on $SO(n, \mathbb{C})$ or on $SO(n)$; ρ is of valence 2 on $SO(n, \mathbb{C})$ or on $SO(n)$.

(2) When n is even: $n = 2r$. Since $C(Q)$ is isomorphic to a total matrix algebra of degree 2^r over \mathbb{C} , $C(Q)$ has a unique (up to equivalence) irreducible representation $\tilde{\rho}$, which is of degree 2^r . The restriction of $\tilde{\rho}$ on $Spin(n, \mathbb{C})$ (on $Spin(n)$) defines a representation ρ of degree 2^r of $Spin(n, \mathbb{C})$ (of $Spin(n)$); ρ is called the spin representation of the group $Spin(n, \mathbb{C})$ (of $Spin(n)$). This representation ρ is, however, not irreducible; ρ is decomposed into the direct sum of two irreducible representations ρ_+ and ρ_- . They are not equivalent to each other, and both are of degree 2^{r-1} . By taking a suitable minimal left ideal L of $C(Q)$ as the representation space of the representation $\tilde{\rho}$, we obtain the representation spaces L^+, L^- of ρ_+, ρ_- , respectively, by putting $L^+ = L \cap C^+(Q)$ and $L^- = L \cap C^-(Q)$. The representation ρ^+ (or ρ^-) is called the **half-spin representation** of the group $Spin(n, \mathbb{C})$ or of the group $Spin(n)$. The elements in the representation space of ρ^+ (or ρ^-) are called **half-spinors**. Again, ρ^+ and ρ^- are not well defined on $SO(n, \mathbb{C})$ or on $SO(n)$.

They are of valence 2 on these groups. The representations of the Lie algebra of the Lie group $SO(n, \mathbb{C})$ (note that this Lie algebra is a \dagger complex simple Lie algebra of type D_r) associated with ρ^+, ρ^- are also called **half-spin representations** of this Lie algebra.

References

- [1] N. Bourbaki, *Éléments de mathématique*, Algèbre, ch. 9, Actuelles Sci. Ind., 1272a, Hermann, 1959.
- [2] C. Chevalley, *The algebraic theory of spinors*, Columbia Univ. Press, 1954.
- [3] M. Eichler, *Quadratische Formen und orthogonale Gruppen*, Springer, 1952.

62 (XI.9) Cluster Sets

A. Cluster Sets of Functions Meromorphic in an Arbitrary Domain

Let D be an arbitrary \dagger domain in the complex z -plane, Γ its boundary, and E a \dagger totally disconnected closed set contained in Γ . Let $w = f(z)$ be a single-valued \dagger meromorphic function defined in D . Then for each point z_0 in Γ , we can define the following sets related to the mapping $w = f(z)$ in the complex w -plane (or on the complex w -sphere $\hat{\mathbb{C}}$).

The Cluster Set. A value α is called a **cluster value** of $f(z)$ at z_0 if there exists a sequence of points $\{z_n\}$ such that

$$z_n \in D, \quad z_n \rightarrow z_0, \quad f(z_n) \rightarrow \alpha.$$

The totality $C_D(f, z_0)$ of all the cluster values of $f(z)$ at z_0 is called the **cluster set** of $f(z)$ at z_0 or, more precisely, the **interior cluster set**. It is a nonempty, closed, but not necessarily connected set.

The Boundary Cluster Set. The set of all values α such that there exist a sequence of points $\{\zeta_n\}$ of $\Gamma - \{z_0\}$ (resp. $\Gamma - \{z_0\} - E$) and a sequence of points $\{w_n\}$ in the complex w -plane satisfying

$$\zeta_n \rightarrow z_0, \quad w_n \in C_D(f, \zeta_n), \quad w_n \rightarrow \alpha$$

is called the **boundary cluster set** of $f(z)$ at z_0 and is denoted by $C_\Gamma(f, z_0)$ (resp. $C_{\Gamma-E}(f, z_0)$). These are closed sets, and

$$C_{\Gamma-E}(f, z_0) \subset C_\Gamma(f, z_0) \subset C_D(f, z_0).$$

If $z_0 \in \Gamma - E$ or z_0 is an isolated point of E , then $C_{\Gamma-E}(f, z_0) = C_\Gamma(f, z_0)$. Furthermore, $C_\Gamma(f, z_0)$

(resp. $C_{\Gamma-E}(f, z_0)$) is empty if and only if z_0 is an isolated boundary point (resp. z_0 is an exterior point of $\Gamma - E$).

Range of Values. The set of values α such that

$$z_n \in D, \quad z_n \rightarrow z_0, \quad f(z_n) = \alpha$$

is called the **range of values** of $f(z)$ at z_0 and is denoted by $R_D(f, z_0)$. In other words, $R_D(f, z_0)$ is the set of values α assumed by $f(z)$ infinitely often in any neighborhood of z_0 and is a $\dagger G_\delta$ -set.

The Asymptotic Set. Let z_0 be an \dagger accessible boundary point of D . If $f(z)$ converges to a value α as z tends to z_0 along a simple arc in D terminating at z_0 , then α is called an **asymptotic value** of f at z_0 . The totality $A_D(f, z_0)$ of asymptotic values of f at z_0 is called the **asymptotic set** of f at z_0 . If z_0 is an inaccessible boundary point, we let $A_D(f, z_0)$ be the empty set.

B. Iversen-Beurling-Kunugui Theorems

Suppose now that E is empty, and put

$$\Omega = C_D(f, z_0) - C_\Gamma(f, z_0).$$

\dagger Iversen's theorems in the case where D is the unit disk $|z| < 1$ and z_0 is a point on $|z| = 1$ are generalized as follows.

(1) $\Omega - R_D(f, z_0) \subset A_D(f, z_0)$ (K. Noshiro, 1936).

(2) First Beurling-Kunugui theorem: If z_0 is not an isolated boundary point, then Ω is an open set.

(3) Second Beurling-Kunugui theorem: Suppose that the open set Ω is not empty. Then $f(z)$ assumes every value belonging to each component Ω_n of Ω infinitely often, with two possible exceptions belonging to Ω_n , that is, $\Omega_n \cap (\hat{C} - R_D(f, z_0))$ consists of at most two values (an extension of \dagger Picard's theorem on an isolated essential singularity).

Next suppose that E is not empty and of \dagger logarithmic capacity zero, and put $\Omega = C_D(f, z_0) - C_{\Gamma-E}(f, z_0)$. If $\alpha \in \Omega - R_D(f, z_0)$, then either α is an asymptotic value of $f(z)$ at z_0 or there exists a sequence of points $\zeta_n \in E$ ($n = 1, 2, \dots$) converging to z_0 such that α is an asymptotic value of $f(z)$ at each ζ_n (Noshiro, 1937). Furthermore, if z_0 is contained in the closure of $\Gamma - E$, then Ω is an open set (which may be empty), and $\Omega - R_D(f, z_0)$ is at most of logarithmic capacity zero (M. Tsuji, 1943). If E is contained in a single component Γ_0 of the boundary Γ , z_0 is contained in the closure of $\Gamma - E$, and Ω is nonempty, then $w = f(z)$ assumes every value belonging to each component Ω_n of Ω infinitely often, with two possible exceptions belonging to Ω_n (Noshiro, 1950). In par-

ticular, if f is bounded in a neighborhood of z_0 , the number of such exceptional values is at most 1. This is still true if each point of E is contained in a boundary component that is a \dagger continuum containing at least two points and Ω is not empty (M. Hervé, 1955). However, this conclusion does not hold if we remove the hypothesis on the set E (K. Matsumoto, 1960).

C. Cluster Sets of Functions Meromorphic in the Unit Disk

Let D be the unit disk $\{|z| < 1\}$, $z_0 = e^{i\theta_0}$ be a fixed point on the unit circumference Γ , A be an open arc of Γ containing z_0 , and E be a set of \dagger linear measure zero such that $z_0 \in E \subset A$. With every $e^{i\theta} \in A - E$ we associate an arbitrary simple arc Λ_θ in D terminating at $e^{i\theta}$ and the **curvilinear cluster set** $C_{\Lambda_\theta}(f, e^{i\theta})$, defined as the set of all values α such that $z_n \in \Lambda_\theta$, $z_n \rightarrow e^{i\theta}$, $f(z_n) \rightarrow \alpha$. We put

$$C_{\Gamma-E}^*(f, z_0) = \bigcap_{r>0} M_r,$$

where M_r denotes the closure of the union $C_{\Lambda_\theta}(f, e^{i\theta})$ for all $e^{i\theta}$ in the intersection of $A - E$ with $|z - z_0| < r$.

By using the above cluster set $C_{\Gamma-E}^*(f, z_0)$ instead of the boundary cluster set $C_{\Gamma-E}(f, z_0)$, we obtain results similar to those in the preceding section (M. Ohtsuka, 1950; Noshiro, 1955). Many interesting results have been obtained by F. Bagemihl and W. Seidel, E. F. Collingwood, O. Lehto and K. I. Virtanen, and others concerning the cluster sets of functions meromorphic in the unit disk. They studied functions of Seidel's class U , **normal** meromorphic functions, and other functions where the class U is the totality of regular bounded functions in the unit disk possessing almost everywhere on $|z| = 1$ radial limits of the constant modulus 1; a nonconstant meromorphic function $f(z)$ in $|z| < 1$ is normal if the family $\{f(T(z))\}$ is \dagger normal in the sense of Montel, where $T(z)$ is an arbitrary conformal mapping of $|z| < 1$ onto itself.

D. The More General Case

The definitions of cluster sets are also available for arbitrary functions for which neither analyticity nor continuity is assumed. If there exist two simple arcs Λ_1 and Λ_2 in the unit disk D terminating at a point $z = e^{i\theta}$ such that

$$C_{\Lambda_1}(f, e^{i\theta}) \cap C_{\Lambda_2}(f, e^{i\theta}) = \emptyset,$$

then $z = e^{i\theta}$ is called an **ambiguous point** of f . Bagemihl proved the following: The set of ambiguous points of an arbitrary complex-valued function defined in the unit disk D is at

most countable [5]. Under the same hypothesis, the set of points $e^{i\theta}$ such that $C_D(f, e^{i\theta}) \neq C_T(f, e^{i\theta})$ is at most countable (Collingwood, 1960). This result shows the importance of introducing the cluster set $C_{T-E}^*(f, z_0)$ that was previously mentioned.

E. History

The theory of cluster sets originated from the \dagger value distribution theory of analytic functions in the neighborhood of their essential singularities. The first systematic results were those of F. Iversen and W. Gross, obtained about 1920. Subsequent significant contributions were made by Seidel, J. L. Doob, M. L. Cartwright, A. Beurling, and others. Since 1940, some important results have been obtained by K. Kunugui, S. Irie, Y. Tôki, Y. Tumura, S. Kametani, Tsuji, Noshiro, and other Japanese mathematicians. Many results have been extended to \dagger pseudoanalytic functions. As can be seen from the Bagemihl ambiguous point theorem, some properties of cluster sets are not intrinsic to analytic mappings [2, 5]. On the other hand, it seems to be an interesting problem to extend the theory of cluster sets to the case of analytic mappings between open Riemann surfaces.

References

- [1] K. Noshiro, Cluster sets, *Erg. Math.*, Springer, 1960.
- [2] E. F. Collingwood and A. J. Lohwater, *The theory of cluster sets*, Cambridge Univ. Press, 1966.
- [3] K. Kunugui, Sur un théorème de MM. Seidel-Beurling, *Proc. Japan Acad.*, 15 (1939), 27–32; Sur un problème de M. A. Beurling, *Proc. Japan Acad.*, 16 (1940), 361–366.
- [4] M. Ohtsuka, On the cluster sets of analytic functions in a Jordan domain, *J. Math. Soc. Japan*, 2 (1950), 1–15.
- [5] F. Bagemihl, Curvilinear cluster sets of arbitrary functions, *Proc. Nat. Acad. Sci. US*, 41 (1955), 379–382.

63 (XVI.3) Coding Theory

A. General Notions of Coding Theory

When we wish to store, to search for, or to send information in the presence of noise efficiently and with the least error, we can

apply various bounds to the efficiency (\rightarrow 213 Information theory). Sophisticated coding operations are required in order to achieve efficiencies as close as possible to the bounds. In the abstract sense, the information is generally understood to be a choice of an element from a finite set X . For implementation, we take a set K of q elements called **alphabets**. Each element of K is called a **letter**. We consider the direct product K^n , i.e., the set of all sequences of n letters. An injection ψ from X into K^n is called **encoding**. The sequence $\psi(x)$ for $x \in X$ is called a **code word**, and the image $\psi(X)$ (a collection of code words) is called a **code**. Such a code is also called a **block code**. The noise is represented by a mapping $\omega: K^n \rightarrow K^n$, but usually it is restricted to a certain subset Ω of the set of such mappings with special properties. For example, we usually assume that the sequences $x \in K^n$ and $\omega(x)$ are different only at less than d letters, where d is a preassigned constant. The inverse mapping φ , i.e., a mapping $\varphi: K^n \rightarrow X$ satisfying $\varphi \circ \psi(x) = x$ for all $x \in X$, is called the **decoding** of ψ . A code ψ satisfying the property $\omega \circ \psi(x) \notin \psi(X)$ for all $x \in X$ and for all $\omega \in \Omega$ is called **error-detecting** with respect to the noise Ω . If ψ has the decoding φ satisfying $\varphi \circ \omega \circ \psi(x) = x$ for all $x \in X$ and for all $\omega \in \Omega$, ψ is called **error-correcting** with respect to the noise Ω . To discuss such properties, we can assume $X = \psi(X) (\subset K^n)$ without restricting the generality, so hereafter we assume this condition. Also, we take $q = 2$ unless explicitly stated otherwise. Because in many communication systems $q = 2$ is commonly adopted, and the generalization for other prime powers q can be obtained naturally.

B. Bounds for the Size of Codes

Let $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in K^n$. The **Hamming distance** $d(x, y) = d_H(x, y)$ between the elements x and y is the total number of unequal bits ($x_i \neq y_i$). We put $d_{\min} = \min\{d_H(x, y) \mid x, y \in X, x \neq y\}$. When $e = \max\{d_H(x, \omega(x)) \mid \omega \in \Omega, x \in X\}$ is less than d_{\min} , the errors due to Ω can be detected, and if $d_{\min} \geq 2e + 1$, the errors can be corrected. The maximal integer t satisfying $d_{\min} \geq 2t + 1$ is often called the **error-correcting capability**. Several important relations are known among d_{\min} , t , n (the length of the code), and $|X|$ (the size of the code word) as follows. **Hamming bound**: $|X| \leq 2^n / \sum_{i=0}^t \binom{n}{i}$. A code satisfying the equality here is called a **perfect code** [6]. **Plotkin bound**: $d_{\min} \leq n|X| / (2(|X| - 1))$. On the other hand, if the **Varsharmov-Gilbert-Sacks bound** $2^{n-k} > \sum_{i=0}^{d_{\min}-2} \binom{n-1}{i}$ is satisfied, there exists an (n, k) -linear code (Section C).

C. Linear Codes

Let $K = GF(2) = \{0, 1\}$, K^n be an n -dimensional vector space over K , and X be a k -dimensional linear subspace. Then X is called a **group code** or (n, k) -**linear code**. In the present case, we can take a suitable basis of K for which there exists a $k \times (n-k)$ matrix P over K such that every vector $\mathbf{x} = (x_1, \dots, x_n)$ representing a code word is expressed as $\mathbf{x} = (\mathbf{z}P, \mathbf{z})$ by a suitable vector \mathbf{z} , and conversely, every vector of this form is a code word, i.e., a vector $\mathbf{x} \in X$ is a code word if and only if

$$(x_1, \dots, x_{n-k}) = (x_{n-k+1}, \dots, x_n)P = \mathbf{0}. \quad (*)$$

Introducing the matrix $H = [I_{n-k}, -P^T]$, where I_{n-k} is the unit matrix of order $n-k$ and T indicates the transpose, the condition $(*)$ is equivalent to $\mathbf{x}H^T = \mathbf{0}$. Therefore, H is called the **parity check matrix**, x_{n-k+1}, \dots, x_n are called the **information bits**, and x_1, \dots, x_{n-k} are called the **check bits**. If \mathbf{x} is deformed to $\mathbf{y} = \omega\mathbf{x}$ by a noise ω , i.e., the original signal \mathbf{x} is transmitted as signal \mathbf{y} , we call $\mathbf{s} = \mathbf{y}H^T = (\mathbf{y} - \mathbf{x})H^T$ the **syndrome** of the transmitted signal \mathbf{y} . If we have an algorithm to determine the error vector $\mathbf{e} = \mathbf{y} - \mathbf{x}$ from the syndrome \mathbf{s} , we have a decoding with error-correcting property. For the linear code, it is evident that $d_{\min} = \min \{ \text{number of nonzero elements of } \mathbf{x} \mid \mathbf{x} (\neq \mathbf{0}) \in X \}$.

The **Hamming code** is given by $n = 2^m - 1$ (m being an integer ≥ 2), $k = n - m$, where the (i, j) -component h_{ij} ($i = 1, \dots, m; j = 1, \dots, n$) of the parity check matrix H is given by the i th bit of the number $j-1$ expressed in the binary (2-adic) number system. The check bits are x_j ($j = 2^0, 2^1, \dots, 2^{m-1}$), and all other bits are the information bits. This code has the error-correcting property with respect to the noise ω for which the nonzero component of $\mathbf{e} = \omega\mathbf{x} - \mathbf{x}$ is at most 1. In fact, from the syndrome $\mathbf{s} = (s_1, \dots, s_m)$, we compute $\hat{j} = s_1 + 2s_2 + \dots + 2^{m-1}s_m$ and put $e_j = 1, e_j = 0$ for all $j \neq \hat{j}$; if $\hat{j} = 0$, we can put $\mathbf{e} = \mathbf{0}$. The Hamming code is a perfect code.

D. Cyclic Codes

The cyclic code is a special case of a linear code. This is the object of one of the most important applications of the theory of finite fields. By using this theory, we can actually construct the codes with high for error-correcting capability in which the encoding and the decoding operations are performed algebraically.

An (n, k) -linear code X over $GF(2)$ is called a **cyclic code** if $\mathbf{x} = (x_1, x_2, \dots, x_n) \in X$ implies $(x_n, x_1, x_2, \dots, x_{n-1}), (x_{n-1}, x_n, x_1, \dots, x_{n-2}), \dots, (x_2,$

$\dots, x_n, x_1) \in X$. To each $\mathbf{x} \in X$, we can associate a polynomial of one variable $x(\xi) = x_1 + x_2\xi + \dots + x_n\xi^{n-1}$ over $GF(2)$. Let $g_X(\xi)$ be the polynomial of the least degree $n-k$ among the polynomials corresponding to $\mathbf{x} \in X$. $g_X(\xi)$ is called the **generator** of X , because the property $\mathbf{x} \in X$ and the property that $x(\xi)$ is divisible by $g_X(\xi)$ are equivalent. $g_X(\xi)$ also divides $\xi^n - 1$. Let the quotient be $h_X(\xi) = (\xi^n - 1)/g_X(\xi)$. Then $\mathbf{x} \in X$ is equivalent to $x(\xi)h_X(\xi) \equiv 0 \pmod{(\xi^n - 1)}$. $h_X(\xi)$ is called the **parity check polynomial**. The **BCH code (Bose-Chaudhuri-Hocquenghem code)** defined below is a typical example of a cyclic code.

Let α be an element of order n in $GF(2^m)$ (the finite field with 2^m elements), and let $g_i(\xi)$ be the polynomial of the lowest degree in $GF(2)[\xi]$ for which α^i is a root ($i = 1, 3, \dots, 2t-1$). The BCH code is given by the least common multiple $g(\xi)$ of the polynomials $g_1(\xi), g_3(\xi), \dots, g_{2t-1}(\xi)$. The BCH code has at least $n - mt$ information bits and satisfies $d_{\min} \geq 2t + 1$, and algebraic decoding methods are known [3, 5]. The Hamming code is a BCH code when $t = 1, n = 2^m - 1$.

E. Other Codes

Other important types of codes include **convolutional codes** (not block codes) for correcting errors in consecutive digits (**burst errors**) and **Goppa codes** (an extension of the BCH class). For correcting burst errors, special cyclic codes or their mixtures are also used. This field of research is closely connected to information theory, algebra, and various applications of combinatorial analysis, such as experimental design [3].

References

- [1] J. Wolfowitz, Coding theorems of information theory, Springer, 1961.
- [2] R. G. Gallager, Low-density parity-check codes, MIT Press, 1963.
- [3] E. R. Berlekamp, Algebraic coding theory, McGraw-Hill, 1968.
- [4] S. Lin, An introduction to error-correcting codes, Prentice-Hall, 1970.
- [5] W. W. Peterson and E. J. Weldon, Jr., Error-correcting codes, MIT Press, second edition, 1972.
- [6] A. Tietäväinen, On the non-existence of perfect codes over finite fields, SIAM J. Appl. Math., 24 (1973), 88–96.
- [7] E. R. Berlekamp, Goppa codes, IEEE Trans. on Information Theory, IT-19 (1973), 590–592.

64 (IX.8) Cohomology Operations

A. General Remarks

The notion of cohomology operations was introduced by L. S. Pontryagin and N. E. Steenrod in order to solve homotopy classification problems (\rightarrow 202 Homotopy Theory). Since then, numerous works have proved the importance of cohomology operations as applied to homotopy theory, differential topology, and other branches of topology. In fact, the use of cohomology operations is indispensable in studying problems related to homotopy groups, characteristic classes of manifolds, etc.

We denote by the symbol $H^*(X; A) = \Sigma H^n(X; A)$ the singular cohomology ring of a topological space X with coefficients in an Abelian group A .

B. Primary Cohomology Operations

A (primary) cohomology operation (or simply an operation) φ is a natural transformation

$$\varphi: \prod_{\lambda} H^{\lambda_1}(; A_{\lambda}) \rightarrow \prod_{\mu} H^{\mu_1}(; B_{\mu})$$

between the cohomology functors defined on the category of topological spaces and continuous mappings. That is, φ is a family of mappings satisfying the following conditions:

(I) For each space X , φ defines a mapping

$$\varphi: \prod_{\lambda} H^{\lambda_1}(X; A_{\lambda}) \rightarrow \prod_{\mu} H^{\mu_1}(X; B_{\mu})$$

that is not necessarily additive.

(2) For each mapping $f: X \rightarrow Y$, the commutativity $f^* \circ \varphi = \varphi \circ f^*$ holds in the diagram

$$\begin{array}{ccc} \prod_{\lambda} H^{\lambda_1}(X; A_{\lambda}) & \xrightarrow{\varphi} & \prod_{\mu} H^{\mu_1}(X; B_{\mu}) \\ \uparrow f^* & & \uparrow f^* \\ \prod_{\lambda} H^{\lambda_1}(Y; A_{\lambda}) & \xrightarrow{\varphi} & \prod_{\mu} H^{\mu_1}(Y; B_{\mu}) \end{array}$$

We list here two trivial examples.

(I) Addition of cohomology groups determines an operation $\varphi: H^{\lambda_1}(X; A) \times H^{\lambda_2}(X; A) \rightarrow H^{\lambda_1 + \lambda_2}(X; A)$.

(II) The cup product determines an operation

$$\varphi: H^{\lambda_1}(X; A_1) \times H^{\lambda_2}(X; A_2) \rightarrow H^{\lambda_1 + \lambda_2}(X; A_1 \otimes A_2)$$

denoted by

$$\varphi(\alpha, \beta) = \alpha \smile \beta.$$

The composite of two cohomology operations is defined in the obvious way. Among cohomology operations the most important ones are operations of one variable. A co-

homology operation of type $(A, l; B, m)$ is a natural transformation

$$\varphi: H^l(; A) \rightarrow H^m(; B).$$

$\mathfrak{D}(A, l; B, m)$ denotes the Abelian group consisting of all such operations.

Denote by $H^m(A, l; B)$ the m th cohomology group of an Eilenberg-MacLane space $K(A, l)$, and let $u \in H^l(A, l; A)$ be the fundamental class of $K(A, l)$. If X is a CW complex, by assigning f^*u to each $f: X \rightarrow K(A, l)$, we obtain a one-to-one correspondence between the set of the homotopy classes $\pi(X, K(A, l))$ and the cohomology group $H^l(X; A)$ (\rightarrow 70 Complexes F). Hence by utilizing condition (2), it can be shown that the value of φ on $H^l(X; A)$ is uniquely determined by its operation on $H^l(A, l; A)$. Thus the assignment $\varphi \rightarrow \varphi u$ defines the isomorphism $\mathfrak{D}(A, l; B, m) \cong H^m(A, l; B)$. Here we should remark that the isomorphism $\pi(K(A, l), K(B, m)) \cong H^m(A, l; B)$ defines a one-to-one correspondence $\mathfrak{D}(A, l; B, m) \rightarrow \pi(K(A, l), K(B, m))$. In some cases $H^m(A, l; B)$ vanishes, for example, when $0 < m < l$, $l = 1 < m$, $A = \mathbb{Z}$, $l = 2$, $m = 2m' + 1$, $A = \mathbb{Z}$; $l = 2l' < m = 2m' + 1$, $A = \mathbb{Z} \cdot B = \mathbb{Q}$; $l = 2l' + 1 < m$, $A = \mathbb{Z}$, $B = \mathbb{Q}$; A is finite, $B = \mathbb{Q}$; etc.

The following four types of operations together with the two above are called elementary operations:

(III) Homomorphisms induced by a coefficient homomorphism: There are homomorphisms $\eta_*: H^l(X; A) \rightarrow H^l(X; B)$ induced by a homomorphism $\eta: A \rightarrow B$.

(IV) Bockstein (cohomology) operation: This operation is given by the connecting homomorphisms $\delta^*: H^l(X; C) \rightarrow H^{l+1}(X; A)$ associated with a short exact sequence $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ of coefficient groups (\rightarrow 200 Homological Algebra). For example, the coefficient sequence $0 \rightarrow \mathbb{Z} \xrightarrow{n} \mathbb{Z} \rightarrow \mathbb{Z}_n \rightarrow 0$ ($\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$) defines a Bockstein operation (or Bockstein homomorphism), which is usually denoted by $(1/n)\delta$ or Δ_n .

(V) Steenrod (or reduced) square operations Sq^i ($i \geq 0$): Sq^i are sequences of operations defined by the following five axioms [2, 5]:

(V1) For each pair of integers $i \geq 0$ and $l \geq 0$,

$$Sq^i: H^l(X; \mathbb{Z}_2) \rightarrow H^{l+i}(X; \mathbb{Z}_2)$$

is a natural transformation of functors that is a homomorphism.

(V2) $Sq^0 = 1$.

(V3) If $\deg x = i$, then $Sq^i x = x \smile x$ (cup product).

(V4) If $\deg x < i$, then $Sq^i x = 0$.

(V5) (Cartan formula)

$$Sq^i(x \smile y) = \sum_{j=0}^i Sq^j x \smile Sq^{i-j} y.$$

These five axioms imply the following two formulas:

(V6) Sq^1 is the Bockstein operation β_2 of the coefficient sequence $0 \rightarrow \mathbf{Z}_2 \rightarrow \mathbf{Z}_4 \rightarrow \mathbf{Z}_2 \rightarrow 0$.

(V7) (**Adem relations**) If $0 < i < 2j$, then

$$Sq^i \circ Sq^j = \sum_{k=0}^{[i/2]} \binom{j-1-k}{i-2k} Sq^{i+j-k} \circ Sq^k.$$

The binomial coefficient is taken mod 2.

We can extend the definition of the Sq^i so that they operate on the relative cohomology groups. Now (V1), (V2), and (V5) imply:

(V8) If $\delta: H^l(Y; \mathbf{Z}_2) \rightarrow H^{l+1}(X, Y; \mathbf{Z}_2)$ is the \dagger coboundary homomorphism, then $\delta \circ Sq^i = Sq^i \circ \delta$.

(V') **Steenrod p th power operations** \mathcal{P}^i ($i \geq 0$): Let p be an odd prime. Then \mathcal{P}^i is a sequence of operations defined by the following five axioms [5].

(V'1) For each pair of integers $i \geq 0$ and $l \geq 0$,

$$\mathcal{P}^i: H^l(X; \mathbf{Z}_p) \rightarrow H^{l+2i(p-1)}(X; \mathbf{Z}_p)$$

is a natural transformation that is a homomorphism.

(V'2) $\mathcal{P}^0 = 1$.

(V'3) If $\deg x = 2k$, then $\mathcal{P}^k x = x^p$.

(V'4) If $\deg x < 2k$, then $\mathcal{P}^k x = 0$.

(V'5) (**Cartan's formula**)

$$\mathcal{P}^i(x \smile y) = \sum_{j=0}^i \mathcal{P}^j x \smile \mathcal{P}^{i-j} y.$$

These axioms imply the **Adem relations** for \mathcal{P}^i and the Bockstein homomorphism β_p associated with the coefficient sequence $0 \rightarrow \mathbf{Z}_p \rightarrow \mathbf{Z}_{p^2} \rightarrow \mathbf{Z}_p \rightarrow 0$ (\rightarrow Appendix A, Table 6.II).

We can extend the definition of \mathcal{P}^i to the relative cohomology groups too, and we obtain $\delta \circ \mathcal{P}^i = \mathcal{P}^i \circ \delta$.

(VI) **Pontryagin p th power operations** \mathfrak{P}_p . Let p be a prime. \mathfrak{P}_p is a system of operations satisfying the following five conditions [3]:

(VI1) For each pair of integers $l \geq 0$ and $h \geq 1$,

$$\mathfrak{P}_p: H^l(X; \mathbf{Z}_{p^h}) \rightarrow H^{pl}(X; \mathbf{Z}_{p^{h+1}})$$

is a natural transformation.

(VI2) If $\eta: \mathbf{Z}_{p^{h+1}} \rightarrow \mathbf{Z}_{p^h}$ is a homomorphism defined by $\eta(1) = 1$, then we have $\eta_* \circ \mathfrak{P}_p x = x^p$.

(VI3) If $p: \mathbf{Z}_{p^h} \rightarrow \mathbf{Z}_{p^{h+1}}$ is a homomorphism defined by $p(1) = p$, then we have

$$\begin{aligned} \mathfrak{P}_p(x + y) \\ = \mathfrak{P}_p x + \mathfrak{P}_p y + \sum_{i=1}^{p-1} \binom{p}{i} p_* (x^i \smile y^{p-i}). \end{aligned}$$

(VI4) $\mathfrak{P}_p(x \smile y) = \mathfrak{P}_p x \smile \mathfrak{P}_p y$.

(VI5) If $p > 2$ and $\deg x = 2k + 1$ (odd), then $\mathfrak{P}_p x = 0$.

Let A, B be finitely generated Abelian groups. Then the computation of $H^*(A, l)$

shows that each element of $\mathfrak{D}(A, l; B, m)$ can be written as the composite of a finite number of the operations of the types (I)–(VI) (H. Cartan).

Here we show some examples. Let u be the fundamental class of $K(\mathbf{Z}_2, 2)$. Then the elements of $H^2(\mathbf{Z}_2, 2; \mathbf{Z}_4) \cong \mathbf{Z}_2$, $H^3(\mathbf{Z}_2, 2; \mathbf{Z}_4) \cong \mathbf{Z}_2$, $H^4(\mathbf{Z}_2, 2; \mathbf{Z}_4) \cong \mathbf{Z}_4$, and $H^5(\mathbf{Z}_2, 2; \mathbf{Z}_4) \cong \mathbf{Z}_4 + \mathbf{Z}_2$ correspond to the cohomology operations $\lambda 2_*$, $\lambda \eta_* \circ (\delta/2)$, $\mu \mathfrak{P}_2$, and $\mu \eta_* \circ (\delta/4) \circ \mathfrak{P}_2 + \lambda 2_* \circ Sq^2 \circ Sq^1$, respectively, where λ, μ denote integers satisfying $0 \leq \lambda \leq 1$, $0 \leq \mu \leq 3$ and $\eta: \mathbf{Z} \rightarrow \mathbf{Z}_4, 2: \mathbf{Z}_2 \rightarrow \mathbf{Z}_4$ stand for the homomorphisms defined by $\eta(1) = 1$, $2(1) = 2$, respectively.

Suppose that we are given sets of integers l_i and m_j . We define the **stable (primary) cohomology operation** φ with respect to these sets of integers l_i and m_j as a system of natural transformations

$$\varphi: \prod_{\lambda} H^{n+l_{\lambda}}(-; A_{\lambda}) \rightarrow \prod_{\mu} H^{n+m_{\mu}}(-; B_{\mu})$$

satisfying the following condition, for all integers $n \geq 0$;

(3) Let $S: H^{l+1}(SX; A) \rightarrow H^l(X; A)$ denote the \dagger suspension. Then the commutativity $S \circ \varphi = \varphi \circ S$ holds in the diagram

$$\begin{array}{ccc} \prod_{\lambda} H^{n+l_{\lambda}}(X; A) & \xrightarrow{\varphi} & \prod_{\mu} H^{n+m_{\mu}}(X; B) \\ \uparrow S & & \uparrow S \\ \prod_{\lambda} H^{n+l_{\lambda}+1}(SX; A) & \xrightarrow{\varphi} & \prod_{\mu} H^{n+m_{\mu}+1}(SX; B). \end{array}$$

This condition is equivalent to the commutativity with the coboundary homomorphisms.

For example, the cohomology operations

$$(-1)^n \beta_p: H^n(X; \mathbf{Z}_p) \rightarrow H^{n+1}(X; \mathbf{Z}_p)$$

define a stable cohomology operation β . Sq^i , \mathfrak{P}^i are also examples of stable cohomology operations.

A stable cohomology operation φ of type (A, B) and of degree q is a sequence of cohomology operations of type $(A, n; B, n+q)$ defined for all integers $n \geq 0$. $\mathfrak{A}(A, B)_q$ denotes the Abelian group consisting of all stable cohomology operations of type (A, B) and of degree q . When $A = B$, $\mathfrak{A}(A) = \sum_{q=0}^{\infty} \mathfrak{A}(A, A)_q$ is a \dagger graded algebra, where multiplication of two operations is given by their composition. Let p be a prime. Then $\mathfrak{A}(\mathbf{Z}_p)$ is called the **Steenrod algebra mod p** and is denoted by $\mathfrak{A}(p)$. $\mathfrak{A}(2)$ is the augmented graded algebra over \mathbf{Z}_2 generated by Sq^i subject to the Adem relations. Suppose that we are given a sequence of non-negative integers $I = (i_1, i_2, \dots, i_k)$. We call I an **admissible sequence** if $i_{s-1} \geq 2i_s$ holds for $2 \leq s \leq k$. We write $Sq^I = Sq^{i_1} \circ Sq^{i_2} \circ \dots \circ Sq^{i_k}$. If I is an admissible sequence, we say that Sq^I is an **admissible monomial**. The admissible monomials form an additive basis for $\mathfrak{A}(2)$, which

has the structure of a \dagger Hopf algebra whose \dagger comultiplication $\psi: \mathfrak{A}(2) \rightarrow \mathfrak{A}(2) \otimes \mathfrak{A}(2)$ is given by $\psi Sq^i = \sum_{j=0}^i Sq^j \otimes Sq^{i-j}$. The dual space $\mathfrak{A}(2)^* = \text{Hom}_{\mathbb{Z}_2}(\mathfrak{A}(2), \mathbb{Z}_2)$ gives a Hopf algebra that is the polynomial algebra generated by ξ_i of degree $2^i - 1$, where ξ_i is the dual element of $Sq^{2^{i-1}} Sq^{2^{i-2}} \dots Sq^1$ with respect to the additive basis. The multiplication $\varphi^*: \mathfrak{A}(2)^* \rightarrow \mathfrak{A}(2)^* \otimes \mathfrak{A}(2)^*$ is given by $\varphi^* \xi_i = \sum_{j=0}^i \xi_{i-j} \otimes \xi_j$.

$\mathfrak{A}(p)$ has properties similar to $\mathfrak{A}(2)$ (p is an odd prime). In particular, $\mathfrak{A}(p)$ is a \dagger Hopf algebra generated by \mathcal{P}^i and β subject to the Adem relations with comultiplication ψ given by $\psi \beta = \beta \otimes 1 + 1 \otimes \beta$ and $\psi \mathcal{P}^i = \sum_{j=0}^i \mathcal{P}^j \otimes \mathcal{P}^{i-j}$. An additive basis of $\mathfrak{A}(p)$ is given by $\{\beta^{\varepsilon_0} \mathcal{P}^{i_1} \beta^{\varepsilon_2} \dots \mathcal{P}^{i_k} \beta^{\varepsilon_k}\}$, where $\varepsilon_i = 0$ or 1 , $i_m \geq p i_{m+1} + \varepsilon_m$ ($m \geq 1$) and $i_k \geq 1$. Such a sequence $I = (\varepsilon_0, i_1, \varepsilon_1, \dots, i_k, \varepsilon_k)$ is also called an **admissible sequence**. Denote the dual element of $\mathcal{P}^{p^{i-1}} \mathcal{P}^{p^{i-2}} \dots \mathcal{P}^1$ and $\mathcal{P}^{p^{i-1}} \mathcal{P}^{p^{i-2}} \dots \mathcal{P}^1 \beta$ by ξ_i and τ_i , respectively; then the dual algebra $\mathfrak{A}(p)^*$ is isomorphic to the tensor product of the polynomial ring over \mathbb{Z}_p generated by ξ_1, ξ_2, \dots and the exterior algebra over \mathbb{Z}_p generated by τ_1, τ_2, \dots . The multiplication φ^* is given by $\varphi^*(\xi_i) = \sum_{j=0}^i \xi_{i-j} \otimes \xi_j$ and $\varphi^*(\tau_i) = \tau_i \otimes 1 + \sum_{j=0}^i \xi_{i-j} \otimes \tau_j$ [5] (\rightarrow Appendix A, Table 6.III).

C. Secondary Cohomology Operations

Here we restrict our attention to a special type of operation treated by J. F. Adams [4] that has been proved to be powerful in applications.

For a specified prime p , we write $H^+(X) = \sum_{i=1}^{\infty} H^i(X; \mathbb{Z}_p)$. We can regard $H^+(X)$ as a \dagger graded left module over $\mathfrak{A}(p)$. Now let C_s ($s=0, 1$) be a pair of \dagger left free modules over $\mathfrak{A}(p)$, and let $d: C_1 \rightarrow C_0$ be a graded homomorphism over $\mathfrak{A}(p)$. Suppose that we are given an element $z \in C_1$ such that $dz=0$. For future purposes, C_0 and C_1 are assumed to have bases $\{c_{0,\lambda}\}$ and $\{c_{1,\mu}\}$ with $\deg c_{0,\lambda} = l_\lambda$ and $\deg c_{1,\mu} = m_\mu$, respectively, in terms of which d has the representation $dc_{1,\mu} = \sum_{\lambda} a_{\mu,\lambda} c_{0,\lambda}$. Then z is expressed in the form $z = \sum_{\mu} b_{\mu} c_{1,\mu}$.

We say that Φ is a **stable secondary cohomology operation** associated with the pair (d, z) if it satisfies the following four axioms:

(1) Let $D^n(d, X)$ be the module consisting of homomorphisms $\varepsilon: C_0 \rightarrow H^+(X)$ over $\mathfrak{A}(p)$ of degree n such that $\varepsilon d = 0$. Putting $\varepsilon(c_{0,\lambda}) = x_\lambda$, we can assume that $D^n(d, X) = \{\varepsilon = \prod_{\lambda} x_\lambda \in \prod_{\lambda} H^{n+l_\lambda}(X) \mid \sum_{\lambda} a_{\mu,\lambda} x_\lambda = 0\}$. Let $Q^n(z, X)$ be the submodule of $H^+(X)$ consisting of elements of the form $\xi(z)$, where ξ is an $\mathfrak{A}(p)$ -homomorphism: $C_1 \rightarrow H^+(X)$ of degree $n-1$.

In other words, $Q^n(z, X) = \sum_{\mu} b_{\mu} H^{n+m_{\mu}-1}(X)$. Then for each $n \geq 0$ and each X , Φ is a mapping

$$\Phi: D^n(d, X) \rightarrow H^+(X)/Q^n(z, X).$$

(2) For each mapping $f: X \rightarrow Y$, the commutativity $f^* \circ \Phi = \Phi \circ f^*$ holds in the diagram

$$\begin{array}{ccc} D^n(d, X) & \xrightarrow{\Phi} & H^+(X)/Q^n(z, X) \\ \uparrow f^* & & \uparrow f^* \\ D^n(d, Y) & \xrightarrow{\Phi} & H^+(Y)/Q^n(z, Y). \end{array}$$

(3) Let $S: H^{n+1}(SX) \rightarrow H^n(X)$ denote the suspension. Then the commutativity $S \circ \Phi = \Phi \circ S$ holds in the diagram

$$\begin{array}{ccc} D^n(d, X) & \xrightarrow{\Phi} & H^+(X)/Q^n(z, X) \\ \uparrow S & & \uparrow S \\ D^{n+1}(d, SX) & \xrightarrow{\Phi} & H^+(SX)/Q^{n+1}(z, SX). \end{array}$$

(4) Let $i: Y \rightarrow X$ be an injection such that $i^* \circ \varepsilon = 0$. That is, $i^* x_\lambda = 0$. We can then find homomorphisms (over $\mathfrak{A}(p)$) $\eta: C_0 \rightarrow H^+(X, Y)$ of degree n and $\zeta: C_1 \rightarrow H^+(Y)$ of degree $n-1$ such that the following diagram commutes:

$$\begin{array}{ccccc} H^+(Y) & \xleftarrow{i^*} & H^+(X) & \xleftarrow{i^*} & H^+(X, Y) & \xleftarrow{\delta^*} & H^+(Y) & \xleftarrow{i^*} & H^+(X) \\ & & \uparrow \eta & & \uparrow \zeta & & & & \\ & & C_0 & \xleftarrow{d} & C_1 & & & & \end{array}$$

Then for any such pair (η, ζ) we have $i^* \circ \Phi \varepsilon \equiv \zeta z \pmod{i^* Q^n(z, X)}$, where δ^* denotes the homomorphism defined by $\delta^* y = (-1)^{\dim y} \delta y$.

For each pair (d, z) , there is at least one associated operation Φ . Existence is proved by means of a \dagger Postnikov system. Let Φ and Φ' be two operations associated with the same (d, z) . Then they differ by a stable primary operation in the sense that there is an element $\varphi \in C_0/dC_1$ such that $\Phi' \varepsilon \equiv \Phi \varepsilon + \varphi \varepsilon \pmod{Q(z, X)}$, where $\varphi \equiv \sum_{\lambda} a_{\lambda} c_{0,\lambda} \pmod{dC_1}$ means $\varphi \varepsilon = \sum_{\lambda} a_{\lambda} x_\lambda$.

For example, define the action of $\mathfrak{A}(2)$ on \mathbb{Z}_2 by the rule $a \cdot v = 0$ if $\deg a > 0$ and $1 \cdot v = v$ for each $v \in \mathbb{Z}_2$. We consider a minimal projective resolution

$$0 \leftarrow \mathbb{Z}_2 \xleftarrow{\varepsilon} C_0 \xleftarrow{d_1} C_1 \xleftarrow{d_2} C_2 \leftarrow \dots$$

(\rightarrow 200 Homological Algebra). First we take $C_0 = \mathfrak{A}(2)$ and define $\varepsilon(1) = 1$. Next we take C_1 to be free over $\mathfrak{A}(2)$ with generators c_i , where $0 \leq i$ and $d_1 c_i = Sq^{2^i}$. Furthermore, we take C_2 to be free over $\mathfrak{A}(2)$ with generators $c_{i,j}$ with $0 \leq i \leq j, j \neq i+1$ and

$$d_2(c_{i,j}) = Sq^{2^j} c_j + \sum_{0 \leq k < j} b_{i,j,k} c_k$$

with $b_{i,j,k} \in \mathfrak{A}(2)$. Here the $d_1 c_i = Sq^{2^i}$ form a minimal set of generators of $\mathfrak{A}(2)$, and the equations $0 = d_1 d_2 c_{i,j} = Sq^{2^j} \circ Sq^{2^j} + \sum_{0 \leq k < j} b_{i,j,k} Sq^{2^k}$ form a minimal set of generators over $\mathfrak{A}(2)$ of the Adem relations. Let $C_1(j)$ be the submodule over $\mathfrak{A}(2)$ generated by c_k with $0 \leq k \leq j$ in C_1 , and let $d_1(j)$ be the restriction of d_1 on $C_1(j)$. We write $z_{i,j} = d_2 c_{i,j}$.

Let $\Phi_{i,j}$ denote an operation associated with $(d_1(j), z_{i,j})$. Then $\Phi_{i,j}$ is defined on the submodule $D_{i,j}^n(X)$ consisting of elements $x \in H^n(X)$ such that $Sq^{2^k}x = 0$ for $0 \leq k \leq j$ and takes values in $H^{n+2^i+2^{j-1}}(X)$ modulo the submodule

$$Q_{i,j}^n(X) = Sq^{2^i}H^{n+2^{j-1}}(X) + \sum_{0 \leq k < j} b_{i,j,k} H^{n+2^k-1}(X).$$

For example, $\Phi_{0,0}: \text{Ker } \beta_2 \rightarrow \text{Cok } \beta_2$ is the generalized Bockstein homomorphism defined by $\delta/2^2$ and $\Phi_{1,1}$ is the operation discovered by J. Adem that appears as the tertiary obstruction of S^n for $n \geq 2$. If $k \geq 3$, we have the relation

$$Sq^{2^{k+1}} \equiv \sum_{\substack{0 \leq i \leq j \leq k \\ j \neq i+1}} a_{i,j,k} \Phi_{i,j} \left(\text{mod } \sum_{\substack{0 \leq i \leq j \leq k \\ j \neq i+1}} a_{i,j,k} Q_{i,j}^n(X) \right).$$

These formulas can be applied to prove the nonexistence of an element with Hopf-invariant 1 in $\pi_{2n-1}(S^n)$ unless $n = 1, 2, 4, 8$ (Adams [4]). Analogous formulas also hold for $\mathfrak{A}(p)$ with $p > 2$. We have no satisfactory theories concerning cohomology operations of orders higher than 2. For cohomology operations in generalized cohomology theory \rightarrow 237 K-Theory.

References

- [1] N. E. Steenrod, Products of cocycles and extension of mappings, *Ann. Math.*, (2) 48 (1947), 290–320.
- [2] J.-P. Serre, Cohomologie modulo 2 des complexes d'Eilenberg-MacLane, *Comment. Math. Helv.*, 27 (1953), 198–232.
- [3] N. E. Steenrod and P. E. Thomas, Cohomology operations derived from cyclic groups, *Comment. Math. Helv.*, 32 (1957), 129–152.
- [4] J. F. Adams, On the non-existence of elements of Hopf invariant one, *Ann. Math.*, (2) 72 (1960), 20–104.
- [5] N. E. Steenrod, Cohomology operations, *Ann. Math. Studies*, Princeton Univ. Press, 1962.
- [6] J. F. Adams, On the structure and applications of the Steenrod algebra, *Comment. Math. Helv.*, 32 (1958), 180–214.
- [7] R. E. Mosher and M. C. Tangora, Cohomology operations and applications in homotopy theory, Harper, 1968.
- [8] E. H. Spanier, Algebraic topology, McGraw-Hill, 1966.
- [9] G. W. Whitehead, Elements of homotopy theory, Springer, 1978.

65 (IX.17) Combinatorial Manifolds

A. Introduction

First of all, we explain some notions in geometric combinatorial topology, also called **PL** (piecewise linear) **topology**, in order to distinguish between the usual topological viewpoint and the combinatorial (or PL) one. By an n -dimensional **simplicial complex** K we mean an n -dimensional, †locally finite and †rectilinear simplicial complex in the Euclidean N -space \mathbf{R}^N , i.e., an n -dimensional †Euclidean simplicial complex (\rightarrow 70 Complexes). The subspace $P = |K|$ of \mathbf{R}^N covered by all simplexes of K is called an n -dimensional **polyhedron**, and K is referred to as a **simplicial division** or simply a **division** of P . A polyhedron covered by a †subcomplex of a division of P is called a subpolyhedron of P . A polyhedron contained in a polyhedron P is a subpolyhedron of P if and only if it is a closed subset of P . An open subset of a polyhedron P is always a polyhedron but not necessarily a subpolyhedron of P [13]. For arbitrary divisions K_1 and K_2 of a polyhedron there is a common †subdivision of K_1 and K_2 . Thus a property of a simplicial complex K which is invariant under subdivision is called a **combinatorial property** of K or of a polyhedron $|K|$. Two polyhedra P and Q are said to be **combinatorially equivalent** if they have †isomorphic divisions K and L , respectively. Then a homeomorphism $f: P \rightarrow Q$ which is induced from an isomorphism $f: K \rightarrow L$ is called a **combinatorial equivalence**. Since a combinatorial equivalence $f: P \rightarrow Q$ is an isomorphism in the category of polyhedra and **PL** (†piecewise linear) **mappings**, (**PL maps**) which is called the **PL category**, it is also called a **PL homeomorphism** or a **PL isomorphism**, and P and Q are said to be **PL homeomorphic** or **PL isomorphic**. A †topological space X which is homeomorphic with an n -dimensional polyhedron P is said to be **triangulable** and is referred to as an n -dimensional topological polyhedron. Then a division K of P , or a homeomorphism $t: P \rightarrow X$, is called a **triangulation** of X . A division K of P is a triangulation of P . However, there is a compact polyhedron of dimension n (≥ 5) which has a triangulation K not combinatorially equivalent to its division (\rightarrow 70 Complexes; for triangulation problems on manifolds \rightarrow Section C below). The notion of a simplicial complex or cell complex had been introduced as a tool to describe topological invariants of triangulable manifolds (H. Poincaré [1]). The notion of an n -dimensional combinatorial manifold (i.e., a

polyhedron locally combinatorially equivalent to an n -dimensional simplex) was established as an important geometric object in combinatorial topology by means of J. H. C. Whitehead's theory of regular neighborhoods [2, 1939, 1940]. The 1960s saw remarkable results in the study of combinatorial manifolds together with the topological study of differentiable manifolds, which gave rise to contemporary combinatorial or PL topology [3, 4]. The fundamental conjecture (Hauptvermutung) for combinatorial manifolds and the combinatorial triangulation problem of manifolds were negatively solved by R. C. Kirby and L. C. Siebenmann [5, 1969]. In the 1970s, important results about general triangulations were obtained by T. A. Chapman [6] R. D. Edwards, and J. W. Cannon [7].

B. Pseudomanifolds and Homology Manifolds

In this section we explain some topological properties of triangulable manifolds which can be defined by means of incidence relations among their triangulations. A n -polyhedron $M = |K|$ is called an n -dimensional **pseudomanifold** if K satisfies the following three conditions: (i) Every simplex of K is either an n -simplex or a face of an n -simplex; (ii) each $(n-1)$ -simplex is a face of at most two n -simplexes; (iii) for any two n -simplexes σ, τ of K , there exists a finite sequence of n -simplexes $\sigma = \sigma_0, \sigma_1, \dots, \sigma_s = \tau$ such that σ_i and σ_{i+1} have an $(n-1)$ -face in common.

Consider the set S of $(n-1)$ -simplexes of K , each of which is a face of only one n -simplex. Then the set of all $\sigma \in S$ together with their faces forms a subcomplex L of K . The polyhedron $|L|$ of L is called the **boundary** of the pseudomanifold $M = |K|$ and is denoted by ∂M as well as by $\partial K = L$. The boundary of a pseudomanifold is not necessarily a pseudomanifold.

Let $|K|$ be an n -dimensional polyhedron (or n -cell complex). A point p in $|K|$ is called a **regular point** of $|K|$ if it has a neighborhood in $|K|$ homeomorphic to an n -dimensional simplex; otherwise, p is called a **singular point**. A pseudomanifold without singular points is a topological manifold, called a **triangulated manifold**. An n -dimensional polyhedron $|K|$ is a pseudomanifold if and only if the set of all regular points in $|K|$ is dense and connected and the set of all the singular points is of dimension less than $n-1$. Thus a connected triangulated manifold is a pseudomanifold whose boundary coincides with the boundary as a topological manifold:

Let τ_1, τ_2 be oriented n -simplexes of an n -dimensional pseudomanifold $M = |K|$. Suppose

that they have an $(n-1)$ -face σ in common. If the incidence numbers of these simplexes with σ satisfy the relation $[\tau_1, \sigma] = -[\tau_2, \sigma]$, then τ_1 and τ_2 are said to be **coherently oriented**. We call $M = |K|$ with all its n -simplexes oriented an **oriented pseudomanifold** if any two n -simplexes τ_1 and τ_2 are coherently oriented whenever they have an $(n-1)$ -face in common. For an oriented pseudomanifold, the formal sum of all its oriented n -simplexes forms an integral cycle of the pair $(M, \partial M)$, which is called the **fundamental cycle** of $(M, \partial M)$. Its homology class, or fundamental class, generates the homology group $H_n(M, \partial M; \mathbb{Z}) = \mathbb{Z}$. When $M = |K|$ is nonorientable, the **fundamental class with coefficient \mathbb{Z}_2** ($\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$) is similarly defined. A pseudomanifold is **orientable** if and only if the topological manifold consisting of all its regular points is orientable. If for each point p of a polyhedron $M = |K|$ the local homology groups are $H_i(p) = 0$ ($i \neq n$) and $H_n(p) = \mathbb{Z}$ or 0 , we call $M = |K|$ an n -dimensional **homology manifold**, where $H_i(p) = H_i(M, M-p; \mathbb{Z})$. If M is connected, it is also a pseudomanifold. The set ∂M of points p of $M = |K|$ with $H_n(p) = 0$ is called the **boundary** of M and is an $(n-1)$ -dimensional homology manifold without boundary.

Let $M = |K|$ be an n -dimensional homology manifold. Given simplexes σ, τ of K , $\sigma < \tau$ means that σ is a face of τ and $\sigma \neq \tau$. By definition, a vertex of the n -barycentric subdivision $Sd K$ of K is the barycenter of a simplex of K , and a k -simplex τ of $Sd K$ is defined to be $|\hat{\sigma}_0 \hat{\sigma}_1 \dots \hat{\sigma}_k|$, where $\sigma_0, \sigma_1, \dots, \sigma_k$ are simplexes of K such that $\sigma_0 < \sigma_1 < \dots < \sigma_k$ and $\hat{\sigma}_i$ is the barycenter of σ_i . In the following we denote $Sd K$ by K' . For a q -simplex σ of K , we denote by $K(\sigma)$ the union of all $(n-q)$ -simplexes $|\hat{\sigma} \hat{\sigma}_1 \dots \hat{\sigma}_{n-q}|$ of K' such that $\sigma < \sigma_1 < \dots < \sigma_{n-q}$. Then $K(\sigma)$ is an $(n-q)$ -dimensional homology manifold and pseudomanifold contained in M . $K(\sigma)$ is called the $(n-q)$ -**dual cell** or simply the **dual** of σ . Let K_σ be the subcomplex of K' consisting of simplexes of K' contained in $K(\sigma)$. Thus we have $|K_\sigma| = |K(\sigma)|$. We denote by $(K(\sigma))'$ the union of simplexes of K_σ which do not intersect with σ . The set of dual cells $K^* = \{K(\sigma) | \sigma \in K\}$ of simplexes of K is called the **dual complex** of K ; and K^* is also called the **dual subdivision** of $|K|$. K^* satisfies the following conditions:

- (i) $\bigcup_{\sigma \in K} K(\sigma) = |K|$. Each $K(\sigma)$ is the n -join $\hat{\sigma} * (K(\sigma))'$ of the barycenter $\hat{\sigma}$ and $(K(\sigma))'$.
- (ii) $\sigma < \tau \Leftrightarrow (K(\sigma))' \supset K(\tau)$ ($\sigma, \tau \in K$), $(K(\sigma))' = \bigcup_{\sigma < \tau \in K} K(\tau)$.
- (iii) If $K(\sigma) \cap K(\tau) \neq \emptyset$, then there exist simplexes of K with σ and τ as faces, and for the least simplex ρ of K with σ and τ as faces, it holds that $K(\sigma) \cap K(\tau) = K(\rho)$.

The boundary $\partial(K(\sigma))$ of the $(n-q)$ -

dimensional homology manifold $K(\sigma)$ satisfies $\partial(K(\sigma)) = (K(\sigma))' \cup ((\partial K)(\sigma))$, where $(\partial K)(\sigma)$ denotes the dual cell of σ in ∂K , and we understand that $(\partial K)(\sigma) = \emptyset$ if $\sigma \notin \partial K$. $K^* \cup (\partial K)^*$ is called the **dual complex** of $(K, \partial K)$ and is denoted by $(K, \partial K)^*$. For an r -simplex σ ($0 \leq r \leq n-1$), it holds that

$$H_i((K(\sigma))', (\partial K(\sigma))'; \mathbb{Z}) = 0, \quad i \neq n-r-1,$$

$$H_{n-r-1}((K(\sigma))', (\partial K(\sigma))'; \mathbb{Z}) \cong \mathbb{Z}.$$

Let $|K^*|^q$ be the union of i -dual cells of $(K, \partial K)^* = K^* \cup (\partial K)^*$ such that $i \leq q$. Let $C_q(K^*) = \tilde{H}_q(|K^*|^q / |K^*|^{q-1})$, and let $\partial_q: C_q(K^*) \rightarrow C_{q-1}(K^*)$ be the \dagger connecting homomorphism $\partial_*: \tilde{H}_q(|K^*|^q / |K^*|^{q-1}) \rightarrow \tilde{H}_{q-1}(|K^*|^{q-1} / |K^*|^{q-2})$ in the \dagger reduced homology exact sequence of $(|K^*|^q, |K^*|^{q-1}, |K^*|^{q-2})$ (\rightarrow 201 Homology Theory F). Then $C_q(K^*)$ is a free Abelian group generated by the q -dual cells of $(K, \partial K)^*$, and $C(K^*) = \{C_q(K^*), \partial_q\}$ is a \dagger chain complex. It follows from the property of $H_*(K(\sigma))', (\partial K(\sigma))'; \mathbb{Z}$ as above that $H_q(C(K^*)) = H_q(|K|)$ as in the case of a \dagger CW complex.

Now suppose that K is oriented. For a q -simplex $\sigma = |a_0 a_1 \dots a_q|$ we give an \dagger orientation $[a_0, a_1, \dots, a_q]$. Furthermore, we choose $a_{q+1}, a_{q+2}, \dots, a_n$ so that $[a_0, a_1, \dots, a_n]$ is the orientation of $|a_0 a_1 \dots a_n|$ induced from the orientation of K , and we give $K(\sigma)$ the orientation determined by $[\hat{\sigma}, \hat{\sigma}_1, \dots, \hat{\sigma}_{n-q}]$, where $\sigma_i = |a_0 a_1 \dots a_{q+i}|$. We define $\sigma \cdot K(\sigma) = 1$ for σ and $K(\sigma)$ thus oriented, and $\tau \cdot K(\sigma) = 0$ if $\tau \neq \sigma$. Let $c = \sum \lambda_i \sigma_i \in C_q(K)$ and $c^* = \sum \mu_j K(\sigma_j)$. Then we define $c \cdot c^* \in \mathbb{Z}$ by $c \cdot c^* = \sum_{i,j} \lambda_i \mu_j (\sigma_i \cdot K(\sigma_j))$. The integer $c \cdot c^*$ is called the **intersection number** of c and c^* . For homology classes $[z] \in H_q(K)$ and $[z^*] \in H_{n-q}(K^*)$, the intersection number $z \cdot z^*$ is independent of the choice of representing cycles, and thus we define the **intersection number** $[z] \cdot [z^*] = z \cdot z^*$.

When X is compact, the isomorphism of the Poincaré-Lefschetz duality theorem $H_q(X, \partial X; \mathbb{R}) \cong H^{n-q}(X; \mathbb{R})$ is given in terms of intersection numbers.

C. Combinatorial Manifolds and PL Manifolds

A polyhedron which is PL homeomorphic with a k -dimensional simplex or its boundary is called a **PL k -ball** or a **PL $(k-1)$ -sphere** and is denoted by B^k, S^{k-1} , respectively. A simplicial complex K or a polyhedron $M = |K|$ is called an n -dimensional **combinatorial manifold** if the \dagger star of each vertex in K is a PL k -ball. If M is an n -dimensional combinatorial manifold, then so is an open subset of M . An n -dimensional PL manifold is a topological manifold X with a distinguished maximal coordinate system $\pi = \{(U_i, h_i)\}$, called a **PL**

structure on X , which consists of homeomorphisms $h_i: U_i \rightarrow V_i$ from open subsets U_i onto open subsets V_i of an n -dimensional simplex such that $h_i \circ h_i^{-1}: h_i(U_i \cap U_j) \rightarrow h_j(U_i \cap U_j)$ are PL homeomorphisms. For a PL manifold (X, π) there is a triangulation $t: M \rightarrow X$ from a combinatorial manifold M such that for each $(U_i, h_i) \in \pi$, $h_i \circ t|_{t^{-1}(U_i)}$ is a PL homeomorphism from an open subset $t^{-1}(U_i)$ of M onto $h_i(U_i)$. Thus the notions of combinatorial manifold and PL manifold are essentially the same. For a \dagger smooth manifold (X, σ) , there is a unique PL homeomorphism class of a combinatorial manifold by means of a \dagger smooth triangulation. In this sense, smooth manifolds are combinatorial manifolds; but the converse of this is not true (\rightarrow 114 Differential Topology C). A triangulation of a topological manifold which is a combinatorial manifold is called a **combinatorial triangulation**. The triangulation problem for n -dimensional topological manifolds, especially the existence and uniqueness problem of their combinatorial triangulations, a long-standing problem, is stated as follows. **CT_n (Combinatorial Triangulation Problem)**: Is every n -dimensional topological manifold homeomorphic with a combinatorial manifold? **CH_n (Hauptvermutung for combinatorial manifolds)**: Homeomorphic n -dimensional combinatorial manifolds should be PL homeomorphic. When $n \leq 3$, CT_n and CH_n hold; this was shown for $n=2$ by T. Radó [15] and for $n=3$ by E. E. Moise [16]. Moreover, in this case, topological manifolds admit unique smooth structures, and homology manifolds are combinatorial manifolds. For $n \geq 5$, \dagger surgery theory was used and developed as an obstruction theory for "homotoping" a homeomorphism $h: M \rightarrow M'$ between n -dimensional combinatorial manifolds to a PL homeomorphism by D. Sullivan (1967). By the discovery of the torus-unfurling method for "isotoping" $h: M \rightarrow M'$ to a PL homeomorphism, Kirby reduced CT_n and CH_n for $n \geq 5$ to CH_n for some specific manifolds, such as thickened tori modulo the boundary, for which surgery methods had been sufficiently developed (for isotopy \rightarrow Section D). The solutions (1)–(9) can be stated as follows.

(1) Classification of combinatorial triangulations. Assume $n \geq 5$. Existence: An n -dimensional closed topological manifold X admits a combinatorial triangulation if and only if an obstruction $\Delta(X) \in H^4(X; \mathbb{Z}_2)$ vanishes. Uniqueness: A homeomorphism $h: M \rightarrow M'$ between n -dimensional closed combinatorial manifolds is isotopic (\rightarrow Section D) to a PL homeomorphism if and only if an obstruction $\Delta(h) \in H^3(X; \mathbb{Z}_2)$ vanishes (Kirby and Siebenmann [5]).

In fact, for each $n \geq 5$, there is an n -

dimensional combinatorial manifold which is homeomorphic to a torus $S^1 \times \dots \times S^1$ but is not PL homeomorphic, and there is an n -dimensional closed manifold which is not homeomorphic to a combinatorial manifold (Siebenmann [18]).

As a special case, the following conjecture is proved for $n \geq 5$.

(2) The n -dimensional **annulus conjecture**.

The closed region bounded by two mutually disjoint locally flat $(n-1)$ -dimensional topological spheres S_1 and S_2 in an n -dimensional sphere S^n is homeomorphic with an n -dimensional annulus $S^{n-1} \times [0, 1]$ (Kirby [19]).

Regarding the uniqueness problem for (not necessarily combinatorial) triangulations of topological manifolds, we mention the following deep result.

(3) **Double suspension theorem** for homology spheres. The double suspensions of homology n -spheres are homeomorphic with S^{n+2} , where a homology n -sphere is a closed topological manifold with the same homology groups as those of S^n (Cannon [7]).

For each $n \geq 3$, there is a homology n -sphere P^n which is a combinatorial manifold with the nontrivial fundamental group (for $n=3$, see Poincaré [1, 1904] and for $n \geq 4$, see M. Newman, 1948). The suspension of P^n is an example of an $(n+1)$ -dimensional homology manifold which is not a topological manifold. The double suspension of P^n is not a combinatorial manifold but is homeomorphic with S^{n+2} . The problem of whether every n -dimensional topological manifold is homeomorphic to a polyhedron is still open for $n \geq 4$. This problem is reduced to a problem in homology 3-spheres (T. Matumoto; D. E. Galewski and R. J. Stern).

In the following we list some results on combinatorial manifolds focusing on Whitehead's theory of regular neighborhoods [2]. Let $M = |K|$ be an n -dimensional combinatorial manifold without boundary. The dual complex K^* consists of PL balls. Thus the division K of M gives rise to a decomposition into PL handles $H(K)$ of M in such a way that for each k -dimensional simplex $\sigma \in K$, the star $St_{K''}(\sigma)$ of its barycenter σ in the second barycentric subdivision K'' of K is a PL handle of index k . If L is a subcomplex of K or K^* , then the star $St_{K''}(|L|)$ of $|L|$ in K'' is a subhandlebody of $H(K)$ or of its dual handlebody $H^*(K) = H(K^*)$, denoted by $H_K(L)$ or $H_{K^*}(L)$, respectively. In particular, if M is closed, orientable, and of dimension 3, then $U = H_K(K^{(1)})$ and $V = H_{K^*}(K_{(1)}^*)$ are PL homeomorphic with a PL boundary-connected sum of some copies of a solid torus $S^1 \times B^2$, and $M = U \cup V$, $U \cap V = \partial U = \partial V$, where $K^{(k)}$ and $K_{(k)}^*$ stand for the k -skele-

ton of K and K^* , respectively. This decomposition $(U, V; \partial U = \partial V)$ of M is called a **Heegaard decomposition** (or **splitting**) of M [12]. We say that a simplicial complex L collapses elementarily to a subcomplex L_1 of L , referred to as an **elementary collapse** $L \searrow L_1$, if $L - L_1 = \{\sigma, \tau\}$, where σ is a k -dimensional simplex of L and τ is a $(k-1)$ -dimensional face of σ which is not a face of any simplex in $L - \{\sigma, \tau\}$. We say that a polyhedron P collapses to a subpolyhedron Q of P or Q expands to P , referred to as a **collapse** $P \searrow Q$ or **expansion** $Q \nearrow P$, if there is a division L of P containing a subcomplex L_1 dividing Q so that there is a sequence of elementary collapsings $L \searrow L_1 \searrow \dots \searrow L_r$. For a subpolyhedron P of a combinatorial manifold M , a subpolyhedron U of M is called a **regular neighborhood** of P in M if (1) U is a combinatorial manifold which is a closed neighborhood of P in M and (2) $U \searrow P$. In general, for a subcomplex L of a simplicial complex K , the star $St_{K''}(|L|)$ of $|L|$ in the second barycentric subdivision K'' of K is called a **second barycentric derived neighborhood** or simply a **derived neighborhood** of $|L|$ in $|K|$ (E. C. Zeeman [4]).

(4) **Regular neighborhood theorem**. Let P be a compact subpolyhedron of a combinatorial manifold M . Existence: Every derived neighborhood of P in M is a regular neighborhood. Uniqueness: Any two regular neighborhoods of P in M are PL homeomorphic by a mapping keeping P pointwise fixed [2, 1939]. On one hand, this theorem is regarded as a combinatorial counterpart to the tubular neighborhood theorem of differential topology. For a topological manifold X with boundary ∂X , there are a neighborhood U of ∂X in X and a homeomorphism $h: \partial X \times [0, 1] \rightarrow U$, called a collar of ∂X in X , such that $h(x, 0) = x$ ($x \in \partial X$). In particular, for a combinatorial manifold $M = X$ we can take U as a derived neighborhood of ∂M in M and h as a PL homeomorphism, called a PL collar of ∂M in M . (For more about normal bundle theory → 147 Fiber Bundles Q). On the other hand, it is also regarded as a combinatorial deformation theorem of a handlebody decomposition $H_{K''}(L)$ of a derived neighborhood $St_{K''}(|L|)$.

Two polyhedra P and Q are said to be **simple homotopy equivalent** if Q is obtained from P by a finite sequence of collapsings and expansions. A composite mapping $f: P \rightarrow Q$ of the inclusion mappings and deformation retractions associated naturally to the expansions and the collapsings from P to Q is called a **simple homotopy equivalence**. Moreover, if P and Q are subpolyhedra of a combinatorial manifold M and those collapsings and expansions take place in M , then their regular neighborhoods are PL homeomorphic [2, 1940].

(5) A simply homotopy equivalence is clear-

ly a homotopy equivalence. Conversely, we have the **simple homotopy theorem**: For the homotopy class of a homotopy equivalence $f: P \rightarrow Q$ between compact connected polyhedra P and Q , there is a well-defined element $W(f)$, called the **Whitehead torsion** of f , in the t Whitehead group $Wh(G)$ of the fundamental group $G = \pi_1(P)$ such that f is homotopic to a simple homotopy equivalence if and only if $W(f) = 0$, ([2, 1950]; \rightarrow also J. W. Milnor [20]).

Let W be a compact combinatorial manifold of dimension $n + 1$. If ∂W consists of two connected components M and N such that the inclusion mappings $M \subset W$ and $N \subset W$ are simple homotopy equivalences, then $(W; M, N)$ is called an **s-cobordism** of dimension $n + 1$. The following theorem of D. Barden, B. Mazur, and J. Stallings is the nonsimply connected PL version of Smale's t h-cobordism theorem.

(6) **s-Cobordism theorem**. For $n \geq 5$, an s-cobordism $(W; M, N)$ of dimension $n + 1$ is PL homeomorphic with $(M \times [0, 1]; M \times 0, M \times 1)$ (B. Mazur [21]; M. Kervaire [22]; C. Weber [23]; J. F. P. Hudson [8]).

(7) Topological invariance of simple homotopy types. Two homeomorphic compact polyhedra are simple homotopy equivalent (T. A. Chapman [6]).

The famous **Poincaré conjecture** that a simply connected closed manifold of dimension 3 is homeomorphic to a 3-sphere is still unsolved despite much effort by many mathematicians. This conjecture is generalized for dimension n as follows.

(8) **Generalized Poincaré conjecture** in dimension n . Any homotopy n -sphere Σ^n is homeomorphic to an n -sphere, where a homotopy n -sphere is defined to be an n -dimensional topological manifold homotopy equivalent to an n -sphere.

This has been affirmatively answered for $n \geq 5$. The first proof was given by S. Smale (\rightarrow 114 Differential Topology F, K). From the viewpoints of collapsing and general position (\rightarrow Section D), J. Stallings [3, 1960] and E. C. Zeeman [11, 198–204] proved the following.

(9) **Engulfing lemma**. For a compact polyhedron P of dimension m in the interior $\mathring{M} = M - \partial M$ of an n -dimensional combinatorial manifold M , there is a PL n -ball in \mathring{M} containing P in its interior (i.e., engulfing P), provided that M is m -connected (i.e., $\pi_i(M) = 0$ ($i \leq m$)) and $n - m \geq 3$. (For the topological version \rightarrow M. H. A. Newman [24].) If a combinatorial manifold $M = |K|$ is a homotopy n -sphere ($n \geq 5$), then by the engulfing lemma, there are PL n -balls B_1 and B_2 engulfing $H_{K^{(n-3)}}$ and $H_{K^{(n-2)}}$ so that $\hat{B}_1 \cup \hat{B}_2 = M$. By the generalized Schoenflies theorem (\rightarrow Section G), $B_2 - \hat{B}_1 = M - \hat{B}_1$ is homeomorphic to

an n -ball. It follows that by the Alexander trick (\rightarrow Section D) M is homeomorphic to an n -sphere. Since CH_n holds for S^n ($n \geq 5$), the homotopy PL n -sphere M is actually PL homeomorphic with S^n .

D. Embeddings of Combinatorial Manifolds

An injective PL mapping $f: P \rightarrow Q$ between polyhedra is called a **PL embedding** if $f(P)$ is a subpolyhedron of Q (\rightarrow Section A). Namely, a PL embedding $f: P \rightarrow Q$ is a PL mapping which is induced from a simplicial injection $f: K \rightarrow L$, called a division of f , for some divisions K and L of P and Q , respectively. In the following we explain PL embeddings from an m -dimensional combinatorial manifold M into an n -dimensional combinatorial manifold N . For a PL embedding $f: M \rightarrow N$, $c = n - m$ is called the **codimension** of f . If $f^{-1}(\partial N) = \partial M$, f is said to be **proper**. Since an injective PL mapping $f: M \rightarrow N$ is a PL embedding if and only if it is proper as a mapping between topological spaces M and N (i.e., the preimages of compact sets are compact), this term “proper” for PL embeddings should not be confused with the term “proper” for mappings. Two PL embeddings $f: M \rightarrow N$ and $f': M' \rightarrow N'$ are said to be (PL) **equivalent** if there is a PL homeomorphism $h: N \rightarrow N'$, called an equivalence from f to f' , such that $h \circ f(M) = f'(M')$. When $M = M'$, f and f' are said to be (PL) **isomorphic** if there is an equivalence $h: N \rightarrow N'$ from f to f' such that $h \circ f = f'$. A **submanifold** of N is defined as a subpolyhedron of N which is a combinatorial manifold. Thus the equivalence class of a PL embedding $f: M \rightarrow N$ is nothing but the PL homeomorphism class of a pair $(N, f(M))$ of N and its submanifold $f(M)$. For a submanifold of N , the same terms as those for the inclusion mapping are used. For every division $f: K \rightarrow L$ of a (proper) PL embedding $f: M \rightarrow N$ and for $x \in M$ we have a (proper) PL embedding $f|_{\text{St}_K(x)}$ from a PL m -ball $\text{St}_K(x)$ into a PL n -ball $\text{St}_L(f(x))$ which represents the t germ of f at $x \in M$. A PL embedding of a PL m -sphere into a PL n -sphere is called a **PL (n, m) -knot** or (n, m) -sphere pair and a proper PL embedding of a PL m -ball into a PL n -ball is called a **PL (n, m) -ball knot** or (n, m) -ball pair. A PL (n, m) -knot (a PL (n, m) -ball knot) is said to be **trivial** or **unknotted** if it is equivalent to the inclusion mapping $\partial I^{m+1} \times 0 \subset \partial I^{n+1}$ ($I^m \subset I^n$), where $I = [-1, 1]$. Then we have the following statements (1)–(5).

(1) **Zeeman's unknotting theorem**. Every PL (n, m) -knot and every PL (n, m) -ball knot are unknotted, provided that $n - m \geq 3$.

In general, an injection $g: X \rightarrow Y$ between

topological manifolds is said to be **locally flat** if for each point $x \in X$, there is a neighborhood U of x in X so that the local germ of f at x is equivalent to the inclusion mapping $U \times 0 \subset U \times I^n$. Then f is necessarily proper, i.e., $f^{-1}(\partial Y) = \partial X$. A PL embedding $f: M \rightarrow N$ is said to be **locally flat** if it is locally flat in the PL category. Namely, $f|_{\text{St}_K(x)}$ is equivalent to $I^m \times 0 \subset I^n$. Conversely, Zeeman's unknotting theorem implies the following.

(2) Every proper PL embedding of codimension ≥ 3 is locally flat.

Since every locally flat PL embedding $f: M \rightarrow N$ admits a normal \dagger block bundle, the classification of $F: M \rightarrow U$ up to isomorphism is reduced to the classification of block bundles (\rightarrow 147 Fiber Bundles Q), where U is a derived neighborhood of $f(M)$ in N . (For the general classification of locally flat PL embeddings by surgical methods \rightarrow C. T. C. Wall [25, Corollary 11.3.1]).

In the codimension 2 case, there exist non-trivial knots (the classical knots) (\rightarrow 235 Knot Theory) and there is a proper PL submanifold which is not locally flat. For example, a complex projective curve V defined by $x^n + y^{n-1}z = 0$ can be regarded as a PL 2-sphere in the complex projective plane CP^2 , and for the inclusion mapping $i: V \subset CP^2$, at $0 = (0:0:1)$, $i|_{\text{St}_V(0)}$ is PL equivalent to a cone of a torus knot $k: S^1 \subset S^3$ of type $(n, n-1)$. For further results in higher-dimensional knot theory (\rightarrow 235 Knot Theory) we mention the works of M. Kato and Y. Matsumoto [26] and S. Capell and J. Shaneson [27].

Here we consider only proper PL embeddings of a fixed m -dimensional combinatorial manifold M into a fixed connected n -dimensional combinatorial manifold N , so that a codimension 0 PL embedding is a PL homeomorphism. Two PL embeddings $f_0, f_1: M \rightarrow N$ are said to be pseudoisotopic if there is a proper PL embedding $F: M \times [0, 1] \rightarrow N \times [0, 1]$, called a pseudoisotopy from f_0 to f_1 , such that $F(x, 0) = (f_0(x), 0)$ and $F(x, 1) = (f_1(x), 1)$ ($x \in M$). Moreover, if F satisfies the level-preserving condition $F(M \times \{t\}) \subset N \times \{t\}$ ($0 \leq t \leq 1$), then f_0 and f_1 are said to be **isotopic**, and F or a homotopy $\{f_t\}$ defined by $F(x, t) = (f_t(x), t)$ ($(x, t) \in M \times [0, 1]$) is called an **isotopy** from f_0 to f_1 . When $M = N$, a PL homeomorphism which is isotopic to the identity, or the isotopy of the identity itself, is called an **ambient isotopy** of N . Two PL embeddings from M to N are said to be **ambient isotopic** if they are isomorphic by an ambient isotopy of N .

Since I^k ($I = [-1, 1]$) is a k -dimensional \dagger convex cell and is a cone $0 * \partial I^k$ of ∂I^k from the center 0, each point x of $I^k - \{0\}$ can be written uniquely as $x = t \cdot u$ for $0 < t \leq 1$ and $u \in \partial I^k$. Thus a PL embedding $f: \partial I^m \rightarrow \partial I^n$ can

be extended to a PL embedding $F: I^m \rightarrow I^n$, called a **cone extension** or simply a **cone** of f , by setting $F(0) = 0$ and $F(t \cdot u) = t \cdot f(u)$ for $t \cdot u \in I^m - \{0\}$. This method of cone extension is often referred to as **Alexander's trick**. By using Alexander's trick and the uniqueness of regular neighborhoods, one can show without much difficulty that (i) the isotopy classes of PL homeomorphisms of PL spheres or PL balls onto themselves are classified by their \dagger degrees ($= \pm 1$) (V. K. A. M. Gugenheim, 1953), and (ii) every PL (n, m) -knot or PL (n, m) -ball knot is equivalent to the standard one $\partial I^{m+1} \times 0 \subset \partial I^{n+1}$ or $I^m \times 0 \subset I^n$ up to isotopy, provided that $n - m \geq 1$. This is why we define the notion of ambient isotopy in a different way from the differentiable category, where the smooth isotopy is always covered by the ambient isotopy. However, for codimension ≥ 3 , because of Zeeman's unknotting theorem, the following holds.

(3) The pseudoisotopy class, the isotopy class, and the ambient isotopy class of a PL embedding of codimension ≥ 3 are the same (J. F. P. Hudson and E. C. Zeeman, 1964).

In order to construct a PL embedding from a continuous mapping within its homotopy class, the notion of general position is useful. A PL mapping $\varphi: M \rightarrow N$ is said to be in **general position** if $\dim \varphi^{-1}(y) \leq 0$ ($y \in N$) and $\dim S(\varphi) \leq 2m - n$, where $S(\varphi)$ is the closure of $\{x \in M \mid \varphi^{-1}(\varphi(x)) \neq \{x\}\}$ in M (J. F. P. Hudson [8]).

(4) **General position theorem.** Every continuous mapping $\psi: M \rightarrow N$ can be approximated by a PL mapping $\varphi: M \rightarrow N$ which is in general position.

The following theorem can be proved by sharpening the general position theorem and the engulfing lemma.

(5) **Irwin's embedding theorem.** Let $\psi: M \rightarrow N$ be a continuous mapping from a compact combinatorial manifold of dimension m into a combinatorial manifold of dimension n such that $\psi|_{\partial M}$ is a PL embedding from ∂M into ∂N . Assume that (i) M is d -connected, (ii) N is $(d+1)$ -connected, and (iii) $n - m \geq 3$, where $d = 2m - n$. Then $\psi: M \rightarrow N$ is homotopic to a proper PL embedding $f: M \rightarrow N$ relative to ∂M . Moreover, if M and N are closed and if (i*) M is $(d+1)$ -connected and (ii*) N is $(d+2)$ -connected, then two homotopic PL embeddings from M into N are ambient isotopic (M. C. Irwin, 1965).

E. 3-Manifolds

There has been a great deal of research done on 3-manifolds, and we have, among others, the following results (1)–(4).

(1) **Sphere theorem.** Let M be an orientable 3-manifold with $\pi_2(M) \neq 0$. Then there exists a PL embedding f of a 2-sphere into M such that f is not homotopic to 0 in M (C. Papakyriakopoulos [28]).

(2) **Dehn's lemma.** Let D be a 2-cell with self-intersections in a 3-manifold M , with its boundary a simple polygon C . If some neighborhood U of C in D has no singularities, then there exists a 2-cell in M without singularities whose boundary is C . In 1910, M. Dehn asserted this lemma, but his proof was incomplete. In 1957, Papakyriakopoulos and T. Homma proved the lemma independently [28, 29].

(3) **Loop theorem.** Let N be the boundary of a 3-manifold M and U be an open set in a component of N . If there is a closed curve in U which is homotopic to 0 in M but not in N , then there exists a simple closed curve in U homotopic to 0 in M but not in N (Papakyriakopoulos [30], J. Stallings [31]).

(4) **Unique decomposition theorem** for a 3-manifold (J. W. Milnor [32], H. Kneser). We assume that all manifolds are connected, oriented, and triangulated 3-manifolds without boundaries and that all homeomorphisms are piecewise linear. Two manifolds M, M' are said to be isomorphic ($M \approx M'$) if there exists an orientation-preserving homeomorphism of M onto M' . Removing an open 3-cell from each of two 3-manifolds M, M' and identifying the boundaries of these removed cells, we obtain a 3-manifold $M \# M'$, called the **connected sum** of M and M' . A manifold that is not isomorphic to the 3-sphere S^3 is called nontrivial. A nontrivial manifold P is called prime if P cannot be decomposed as $P = M_1 \# M_2$ with M_1 and M_2 both nontrivial. A manifold M is called **irreducible** if every 2-sphere in M bounds a 3-cell. Then from the sphere theorem the following results can be deduced: If a compact 3-manifold M is nontrivial, then M is isomorphic to a connected sum $P_1 \# P_2 \# \dots \# P_k$ of prime manifolds P_i ($i = 1, 2, \dots, k$), where P_1, P_2, \dots, P_k are determined uniquely up to order. Every irreducible 3-manifold M has $\pi_2(M) = 0$. Conversely, if a 3-manifold M has $\pi_2(M) = 0$, then M is irreducible, provided that the Poincaré conjecture is correct. (For further results on 3-manifolds \rightarrow F. Waldhausen [10] and J. Hempel [12].)

Irreducible sufficiently large 3-manifolds [10] can further be decomposed into Seifert fibered spaces and "simple manifolds" which contain no essential tori (W. Jaco and P. B. Shalen, K. Johannson). Existence of hyperbolic structures on "simple manifolds" has been extensively studied by W. Thurston.

F. Wild Spaces

Nonclosed 3-manifolds behave very differently from closed 3-manifolds. For example, there exists an open 3-manifold U that has the same homotopy type as an open 3-cell but is not homeomorphic to it (M. H. A. Newman and J. H. C. Whitehead, 1937). The construction of this and similar examples generally involves infinite processes. Such examples, which do not exhibit the properties we expect, are commonly termed **pathological** (or **wild**) spaces. For the manifold U above, the product $U \times \mathbf{R}^1$ is found to be homeomorphic to the Euclidean 4-space \mathbf{R}^4 . R. Moore proved that the quotient space of the Euclidean 2-space \mathbf{R}^2 by any upper semicontinuous decomposition consisting of continua not separating \mathbf{R}^2 is homeomorphic to \mathbf{R}^2 . On the other hand, R. H. Bing (1959) gave an example of a similar upper semicontinuous decomposition of \mathbf{R}^3 for which the quotient space B is not homeomorphic to \mathbf{R}^3 and is not even a manifold. This quotient space has the property $B \times \mathbf{R}^1 \approx \mathbf{R}^4$. Bing (1957) also proved that at most countably many copies of a wild closed surface can simultaneously be embedded in \mathbf{R}^3 . On the other hand, Stallings (1960) proved that there exists a wild disk of which a set of copies with the power of the continuum can simultaneously be embedded in \mathbf{R}^3 [11].

The study of the quotient space of n -dimensional manifolds by cell-like decompositions has been fully generalized by Edwards and Cannon [7] for $n \geq 5$ to prove the double suspension theorem of homology spheres (\rightarrow Section C (3)).

G. The Schoenflies Theorem

In 1906, A. M. Schoenflies proved that every simple closed curve in the plane is the boundary of a 2-cell. This result is sharper than the Jordan curve theorem and is called the **Schoenflies theorem**. However there was a gap in the proof, and the complete proof was given by L. E. J. Brouwer in 1910.

The following higher-dimensional analog of the above theorem is called the **Schoenflies problem**: Is every $(n-1)$ -sphere embedded in the n -dimensional Euclidean space \mathbf{R}^n ($n \geq 3$) the boundary of an embedded n -disk?

For the case where the embedding is topological, L. Antoine constructed a counterexample, called **Antoine's necklace**, for $n = 3$ in 1921 and solved this problem negatively. Also, J. W. Alexander constructed another counterexample, called **Alexander's horned sphere**, in 1924 (Fig. 1). A similar example was given by R. H. Fox and E. Artin in 1948 by making use

of knot theory (Fig. 2). The study of wild embeddings began with these examples, and many interesting results were obtained (Bing).

For the case where the embedding is PL or smooth, the Schoenflies problem has been solved affirmatively for $n=3$ (Alexander).

M. Brown (1960) and B. Mazur (1959) proved the following: Let $h: S^{n-1} \times [-1, 1] \rightarrow S^n$ be a topological embedding; then $S^n - h(S^{n-1} \times \{0\})$ consists of connected open sets D_+ and D_- of S^n , and the closures of D_+ and D_- are both homeomorphic to the n -disk. This result solves the Schoenflies problem affirmatively for $n \geq 5$ when the embedding is PL and locally flat or smooth.

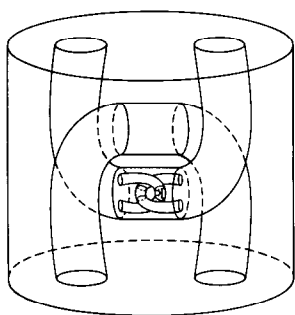


Fig. 1

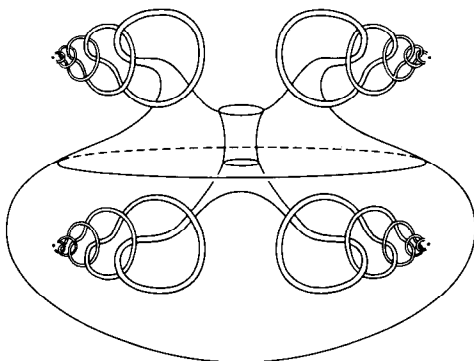


Fig. 2

References

[1] H. Poincaré, *Analysis situs*, J. Ecole Polytech., 1 (1895), 1–121; Cinquième complément à l'analysis situs, Rend. Circ. Mat. Palermo, 18 (1904), 45–110.
[2] J. H. C. Whitehead, *Simplicial spaces, nuclei and m -groups*, Proc. London Math. Soc., 45 (1939), 243–327; *On the homotopy type of manifolds*, Ann. Math., (2) 41 (1940), 825–832; *Simple homotopy types*, Amer. J. Math., 72 (1950), 1–57.
[3] J. Stallings, *Polyhedral homotopy spheres*, Bull. Amer. Math. Soc., 66 (1960), 485–488; *Notes on polyhedral topology*, Tata Inst., 1967.

[4] E. C. Zeeman, *Unknotting combinatorial balls*, Ann. Math., (2) 78 (1963), 501–526; *Seminars on combinatorial topology*, Publ. Math. Inst. HES and Univ. of Warwick (Conventry), 1963–1966.
[5] R. C. Kirby and L. C. Siebenmann, *On the triangulation of manifolds and the Hauptvermutung*, Bull. Amer. Math. Soc., 75 (1969), 742–749; *Foundational essays on topological manifolds, smoothings and triangulations*, Ann. Math. Studies 88, Princeton Univ. Press, 1977.
[6] T. A. Chapman, *Topological invariance of Whitehead torsion*, Amer. J. Math., 96 (1974), 488–497.
[7] J. W. Cannon, *Shrinking cell-like decompositions of manifolds. Codimension three*, Ann. Math., (2) 110 (1979), 83–112.
[8] J. F. P. Hudson, *Piecewise linear topology*, Benjamin, 1969.
[9] C. P. Rourke and B. J. Sanderson, *Introduction to piecewise linear topology*, Erg. Math., 69 (1972).
[10] F. Waldhausen, *On irreducible 3-manifolds which are sufficiently large*, Ann. Math., (2) 87 (1968), 56–88, Amer. Math. Soc. Proc. Symposia in Pure Math., vol. 32, part 2, 1978.
[11] M. K. Fort, Jr. (ed.), *Topology of 3-manifolds and related topics*, Prentice-Hall, 1962.
[12] J. Hempel, *3-manifolds*, Ann. Math. Study 86, Princeton Univ. Press, 1976.
[13] P. Alexandroff and H. Hopf, *Topologie I*, Springer, 1935.
[14] H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Teubner, 1934.
[15] T. Rado, *Ueber den Begriff der Riemannschen Fläche*, Acta Litt. Sci. Szeged, 2 (1925), 101–121.
[16] E. E. Moise, *Affine structures on 3-manifolds*, Ann. Math., (2) 56 (1952), 96–114.
[17] D. Sullivan, *On the Hauptvermutung for manifolds*, Bull. Amer. Math. Soc., 73 (1967), 698–700.
[18] L. C. Siebenmann, *Topological manifolds*, Proc. Int. Congress Math. (Nice, 1970), vol. 2, 133–163.
[19] R. C. Kirby, *Stable homeomorphisms and the annulus conjecture*, Ann. Math., (2) 89 (1969), 575–582.
[20] J. W. Milnor, *Whitehead torsion*, Bull. Amer. Math. Soc., 72 (1966), 358–426.
[21] B. Mazur, *Differential topology from the viewpoint of simple homotopy theory*, Publ. Math. Inst. HES, 15 (1963), 5–93.
[22] M. Kervaire, *Le Théorème de Barden-Mazur-Stallings*, Comment. Math. Helv., 40 (1965), 31–42.
[23] C. Weber, *L'élimination des points doubles dans le cas combinatoire*, Comment. Math. Helv., 41 (1966), 179–182.

- [24] M. H. A. Newman, The engulfing theorem for topological manifolds, *Ann. Math.*, (2) 84 (1966), 555–571.
- [25] C. T. C. Wall, *Surgery on compact manifolds*, Academic Press, 1970.
- [26] M. Kato and Y. Matsumoto, Simply connected surgery of submanifolds in codimension two, *J. Math. Soc. Japan*, 24 (1972), 586–608.
- [27] S. Cappell and J. Shaneson, The codimension two placement problem and homology equivalent manifolds, *Ann. Math.*, (2) 99 (1974), 277–348.
- [28] C. Papakyriakopoulos, On Dehn's lemma and the asphericity of knots, *Ann. Math.*, (2) 66 (1957), 1–26.
- [29] T. Homma, On Dehn's lemma for S^3 , *Yokohama Math. J.*, 5 (1957), 223–244.
- [30] C. Papakyriakopoulos, On solid tori, *Proc. London Math. Soc.*, 7 (1957), 281–299.
- [31] J. Stallings, On the loop theorem, *Ann. Math.*, (2) 72 (1960), 12–19.
- [32] J. W. Milnor, A unique factorization theorem for 3-manifolds, *Amer. J. Math.*, 84 (1962), 1–7.
- [33] W. Thurston, *The geometry and topology of three-manifolds*, Princeton Univ. Press, 1982.

66 (XVI.10) Combinatorics

A. General Review of Combinatorics

Combinatorics is concerned mainly with problems of discrete sets, such as the enumeration of subsets satisfying certain conditions, or their actual construction and the selection of an optimal subset with respect to a suitable criterion. Generally speaking, both the theoretical analysis and the construction of discrete sets are much more difficult than those problems in analysis concerning infinite sets with suitable topologies.

The main emphasis and the name of this field have changed from time to time and from person to person. Other names such as **combinatorial analysis**, **combinatorial theory**, or **combinatorial mathematics** are commonly used and mean almost the same thing. Recently, the expression “**discrete mathematics**” has also been used to describe the same field. In many cases a problem originally conceived as number-theoretic, algebraic, analytic, or geometric has eventually turned out to be combinatorial (\rightarrow 102 Design of Experiments, 241 Latin Squares).

[†]Graph theory, originally a combinatorial study of 1-dimensional topological complexes, has developed into an important branch of mathematics.

In this article we discuss various topics not yet considered to be established branches or organized fields of mathematics. The topics treated under the title “combinatorics” are at present not unified, and consist of a mixture of isolated techniques. However, recent progress with electronic computers is playing an important role in the solution of various combinatorial problems arising from large systems, and the study of combinatorics is now quite active both in theoretical development and application. We now have special journals in this field, such as the *Journal of Combinatorial Theory* (Academic Press), *Combinatorica* (Elsevier), the *European Journal of Combinatorics* (Academic Press), *Discrete Mathematics* (North-Holland), and *Discrete Applied Mathematics* (North-Holland).

B. Three Fundamental Principles

The most fundamental among the various principles in combinatorics are the following three principles for enumeration problems. Let Ω be a finite set and let $|A|$ denote the number of elements in a subset A of Ω .

(1) **Rule of sums:** If $A \cap B = \emptyset$, then $|A + B| = |A| + |B|$.

(2) **Rule of products:** $|A \times B| = |A| \times |B|$, where $A \times B$ is the [†]direct product of the two sets A and B .

(3) **Principle of inclusion and exclusion:**

$$\begin{aligned}
 &|A_1 \cup A_2 \cup \dots \cup A_n| \\
 &= \sum_i |A_i| - \sum_{i < j} |A_i \cap A_j| + \sum_{i < j < k} |A_i \cap A_j \cap A_k| \\
 &\quad - \dots + (-1)^{n-1} |A_1 \cap A_2 \cap \dots \cap A_n|.
 \end{aligned}$$

These principles can be generalized as statements concerning various measures.

C. Möbius Inversion Formula

Let P be an ordered set whose arbitrary interval is a finite set. Let $[P \times P] = \{(x, y) \in P \times P \mid x \leq y\}$ be the set of intervals of P . A function satisfying the following two conditions is determined uniquely and is called the **Möbius function** over P :

- (i) $\mu(x, x) = 1$ for all $x \in P$;
- (ii) $\sum_{x: x \leq z \leq y} \mu(x, z) = 0$ for all $x, y \in P$ with $x < y$ and $x \neq y$.

When K is a field, for a mapping $f: [P \times P] \rightarrow K$, we can define another mapping $g: [P \times P]$

$P \rightarrow K$ by

$$g(x, y) = \sum_{z \in [x, y]} f(x, z).$$

Then we have the **Möbius inversion formula**

$$f(x, y) = \sum_{z \in [x, y]} g(x, z) \mu(z, y).$$

This formula implies, as its special cases, many important formulas, such as the Möbius inversion formula in number theory, the principle of inclusion and exclusion, and the inverse relation of difference and summation.

D. Some Special Sequences

In the enumeration problem, many important sequences $\{a_n\}$ are given by the generating function $f(x) = \sum a_n x^n$. Typical examples are as follows: (1) Binomial coefficients: $(1+x)^m = \sum \binom{m}{n} x^n$. (2) Multinomial coefficients: $(x_1 + \dots + x_r)^m = \sum \binom{m}{n_1, \dots, n_r} x_1^{n_1} \dots x_r^{n_r}$. (3) Multiple combination $a_n = {}_m H_n = \binom{m+n-1}{n} (1-x)^{-m} = \sum \binom{m}{n} H_n x^n$. (4) Bell numbers: $\exp(e^x - 1) = \sum B_n x^n / n!$ (5) The **Stirling number of the second kind** S_m^n is the total number of bijections from N_n to N_m divided by $m!$, where $N_n = \{1, 2, \dots, n\}$. This is equivalent to the total number of partitions of N_n into m nonempty blocks: $(e^x - 1)^m / m! = \sum_{n=m}^{\infty} S_m^n x^n / n!$.

In general, when the generating function $f(x)$ is an analytic function of one or several variables as above, we can obtain recurrence formulas for the sequence from the functional equations or the differential equations which $f(x)$ satisfies; or we can get asymptotic estimates by using the saddle point method.

E. Pólya's Enumeration Theorem

Let U be a finite set, V be a (finite or infinite) set, G be a transformation group operating on U , and w be a function assigning to each element of V a "weight." A weight is an element of a ring and is usually defined by a suitable generating function. The weight of a function $f: U \rightarrow V$ is defined by $w(f) = \prod_{u \in U} w(f(u))$. Two functions f_1 and f_2 are considered identical ($f_1 \equiv f_2$) if there exists a $\pi \in G$ such that $f_1 = f_2 \circ \pi$. Put $F = V^U / (\Xi)$, the set of all equivalence classes of the relation Ξ . $w(\tilde{f})$ is well-defined for $\tilde{f} \in F$, because $f_1 \equiv f_2$ implies $w(f_1) = w(f_2)$. Then we have the equality $\sum_{\tilde{f} \in F} w(\tilde{f}) = P_G(\sum_{v \in V} w(v), \sum_{v \in V} w(v)^2, \dots, \sum_{v \in V} w(v)^m)$. Here, $P_G(y_1, y_2, \dots, y_m)$ is the **cycle index** of G , which is given by $n^{-1} \sum_{\pi \in G} y_1^{j_1(\pi)} y_2^{j_2(\pi)} \dots y_m^{j_m(\pi)}$, where $n = |G|$, $m = |U|$, and $j_k(\pi)$ is the number of cyclic permutations of length k contained in π . This is called **Pólya's enumeration theorem**.

When $G = S_m$ (the symmetric group), $P_{S_m}(y_1, \dots, y_m) = \sum_{\lambda_k \geq 0, \sum k \lambda_k = m} (\lambda_1! 2^{\lambda_2} \lambda_2! \dots m^{\lambda_m} \lambda_m!)^{-1} y_1^{\lambda_1}$

$\dots y_m^{\lambda_m}$. When $G = A_m$ (the alternating group), $P_{A_m}(y_1, \dots, y_m) = \sum_{\lambda_k \geq 0, k \lambda_k = m, \sum \lambda_k = \text{even}} 2 (\lambda_1! 2^{\lambda_2} \lambda_2! \dots m^{\lambda_m} \lambda_m!)^{-1} y_1^{\lambda_1} \dots y_m^{\lambda_m}$. When $G = C_m$ (the cyclic group), $P_{C_m}(y_1, \dots, y_m) = m^{-1} \sum_{k|m} \varphi(k) y_k^{m/k}$, where $\varphi(k)$ is the Euler function, i.e., the number of integers less than k and coprime to k . This theorem is widely applicable to various problems with suitable choices of weights. Several generalizations have also been studied (see, e.g., [3]).

F. The Notion of Polymatroids

The theory of matroids was originally introduced as an abstraction of the notion of linear dependence in a vector space or of some properties of graphs or of the notion of algebraic dependence in the theory of fields (see, e.g., [4, 5]). But recently, as many important applications were discovered, it has become an important branch of combinatorics (see [6–8]).

Let μ be a function from a lattice L to a totally ordered module R . μ is called **submodular** if $\mu(x) + \mu(y) \geq \mu(x \vee y) + \mu(x \wedge y)$ for every pair of elements $x, y \in L$. The set of $a \in L$ such that $\mu(a) = \min_{x \in L} \mu(x)$ constitutes a sublattice in L . Let L be the family of all subsets of a finite set E . The sublattices K in $L = 2^E$ and the pseudo-orders over E correspond in a one-to-one manner as follows (here pseudo-order means a binary relation taken as an ordering without assuming the antisymmetric law): (i) The class of a partition P of E corresponds to the difference of two subsets which are the immediate predecessor and immediate successor in K , or the complement with respect to E of the maximum element in K , or the complement of the minimum element. (ii) The complement of the maximum element in K is maximum in the order of P , and the minimum element in K is minimum in the order of P . (iii) Two classes ξ_1, ξ_2 in P , given as the difference of two elements in K , satisfy the relation $\xi_1 \geq \xi_2$ with respect to the order of P if and only if all elements of K such that $x \supset \xi_1$ in E satisfy $x \supset \xi_2$.

Now let E be a finite set, $L = 2^E$, and let \mathbf{R} be the real number field with the usual addition and order relations. If a submodular function $\mu: 2^E \rightarrow \mathbf{R}$ satisfies the following conditions, then $\mathbf{P} = (E, \mu)$ is called a **polymatroid** over E , and μ is called the **rank function** over \mathbf{P} . (i) $\mu(\emptyset) = 0$, where \emptyset is the empty set. (ii) $x \subset y$ ($\subset E$) implies $\mu(x) \leq \mu(y)$. A function $u: E \rightarrow \mathbf{R}$ is automatically extended to $u: 2^E \rightarrow \mathbf{R}$ by $u(x) = \sum_{e \in x} u(e)$ ($x \subset E$), which is also denoted by the same symbol u . If a function u satisfies $0 \leq u(x) \leq \mu(x)$ for every $x \in E$, then u is called an **independent vector** over \mathbf{P} . An independent

vector u with $u(E) = \mu(E)$ is called a **base** of \mathbf{P} . A function $u: E \rightarrow \mathbf{R}$ is identified with an element in \mathbf{R}^n , where $n = |E|$. Then a set of independent vectors in \mathbf{P} forms a \dagger convex polyhedron in \mathbf{R}^n . If we denote the polyhedron by \mathbf{P} , then such a \mathbf{P} is characterized by the following two conditions: (P1) If $u \in \mathbf{P}$ then each v satisfying $0 \leq v(e) \leq u(e)$ for every $e \in E$ also belongs to \mathbf{P} . (P2) If two $u, v \in \mathbf{P}$ satisfy $u(E) < v(E)$, then there exists $w \in \mathbf{P}$, $w \neq u$, satisfying $u(e) \leq w(e) \leq \max[u(e), v(e)]$ for every $e \in E$. If there are two polymatroids $\mathbf{P}_1 = (E, \mu_1)$ and $\mathbf{P}_2 = (E, \mu_2)$ over the same finite set E , then we have $\max\{u(E) \mid u \in \mathbf{P}_1 \cap \mathbf{P}_2\} = \min\{\mu_1(x) + \mu_2(E - x) \mid x \subseteq E\}$. According to \dagger information theory, an example of a polymatroid (E, μ) is given by the following: $\mu(x)$, $x \subseteq E$, is the simultaneous \dagger entropy of a subset x of the information source E with correlation.

G. Matroids

A polymatroid $\mathbf{P} = (E, \mu)$ is called a **matroid** if the values of μ are always integers and $\mu(x) \leq |x|$ = the cardinality of the set x . A set $x \subseteq E$ is called an **independent set** if $\mu(x) = |x|$, and a **dependent set** otherwise. An independent set with maximal cardinality is called a **base**. A minimal dependent set is called a **circuit**. The set y with $\max\{|y| \mid y \supset x, \mu(x) = \mu(y)\}$ is uniquely determined by x , which is called the **closure** of x and denoted by $\text{cl } x$. Several axioms stated in terms of these notions exist, and can be used to characterize matroids. Some of them are:

(1) The axiom for the family of independent sets \mathcal{I} : (i) $\emptyset \in \mathcal{I}$; (ii) $x \in \mathcal{I}$ and $y \subset x$ implies $y \in \mathcal{I}$; (iii) If $x, y \in \mathcal{I}$ and $|x| < |y|$, then there exists an $e \in E$ in $y - x$ such that $x \cup \{e\} \in \mathcal{I}$.

(2) The axiom for the family of bases \mathcal{B} : (i) \mathcal{B} is not empty; (ii) If $b, b' \in \mathcal{B}$ and $b \neq b'$, then for every $e \in b - b'$, there exists an e' such that $e' \in b' - b$ and $(b - \{e\}) \cup \{e'\} \in \mathcal{B}$.

(3) The axiom for the family of circuits \mathcal{C} : (i) $\emptyset \in \mathcal{C}$; (ii) For every pair $x, y \in \mathcal{C}$, y is not a proper subset of x ; (iii) If $x, y \in \mathcal{C}$ and $x \neq y$, then for every $e \in x \cap y$ there exists a $z \in \mathcal{C}$ such that $z \subset x \cup y - \{e\}$.

(4) The axiom for the closure function cl : $2^E \rightarrow 2^E$: (i) $x \subset \text{cl } x = \text{cl}(\text{cl } x)$ for every $x \subseteq E$; (ii) $x \subset y$ implies $\text{cl } x \subset \text{cl } y$; (iii) If $e \in \text{cl}(x \cup \{e'\}) - \text{cl } x$, then $e' \in \text{cl}(x \cup \{e'\}) - \text{cl } x$.

The following are typical examples of matroids: (1) Let E be the family of finite set of vectors in a vector space V and \mathcal{I} be the family of sets of linearly independent vectors. (2) Let E be the set of edges of a \dagger graph and \mathcal{C} be the set of edges constituting a fundamental circuit. (3) Let E be a finite set and S be a family of subsets of E . $x = \{e_1, \dots, e_r\} \subseteq E$ is

defined to be independent if there exist different elements s_1, \dots, s_r in S such that $e_i \in s_i$ for every $i = 1, \dots, r$.

Let a weight function $w: E \rightarrow \mathbf{R}$ be defined for every element $e \in E$ in a matroid (E, μ) . To obtain a base b of the matroid with minimal weight $w(b) = \sum_{e \in b} w(e)$, we can apply the following procedure, called the **greedy algorithm**: (1) Arrange the elements e_1, \dots, e_n of E in such a manner that $w(e_1) \leq \dots \leq w(e_n)$. (2) Put $b^{(0)} := \emptyset$, and $i := 0$. (3) If $x \equiv b^{(i)} \cup \{e_{i+1}\} \in \mathcal{I}$, then $b^{(i+1)} := x$; otherwise, $b^{(i+1)} := b^{(i)}$. (4) Repeat process (3) for $i = 0, 1, \dots, n-1$. Then the final $b^{(n)}$ is the base of minimal weight.

H. Operations on Matroids

For a matroid $M = (E, \mu)$ with the family \mathcal{B} of its base, $\mathcal{B}^* = \{E - x \mid x \in \mathcal{B}\}$ also satisfies the axioms of the base. The matroid $M^* = (E, \mu^*)$ with \mathcal{B}^* as its base is called the **dual** of M . Here we define $\mu^*(x) = |x| + \mu(E - x) - \mu(E)$. When $F \subset E$, $M|F = (F, \mu_F)$ is called the **reduction** of $M = (E, \mu)$ to F , where μ_F is the restriction of μ on 2^F . A matroid $M \times F = (M^*|F)^*$ is called the **contraction** of M to F . A matroid given by the form $(M|F) \times F'$ is called a **minor** of M . From two matroids $M_1 = (E, \mu_1)$ and $M_2 = (E, \mu_2)$ over the same set E , we can construct a matroid $M = M_1 \vee M_2 = (E, \mu)$ whose rank function $\mu(x)$ is defined by the minimum of $\{\mu_1(y) + \mu_2(y) + |x - y| \mid y \subset x\}$. This is called the **union** of M_1 and M_2 .

Let $M = (E, \mu)$ be a matroid and α be a positive constant. The function $\tilde{\mu}(x) = \mu(x) - \alpha|x|$ is submodular. The set L of x which minimizes $\tilde{\mu}(x)$ is a sublattice of 2^E . The partition of E and the order relation of its classes defined by L are called the **principal partition** of M with respect to the parameter α . When the value of the parameter α is not specified, α is taken to be 2.

As was seen in example (1) at the end of the preceding section, the family of vectors in a vector space V over a field K is a matroid M . A matroid isomorphic to such an M is called **linearly representable** over K . If $K = GF(p)$, the finite field with p elements, then it is called a **p -ary matroid**. A matroid linearly representable over any field is called **regular**. If $[G]$ is a matroid isomorphic to the matroid defined by the family of edges of a graph as shown in example (2), it is called **graphic**. If the dual M^* is graphic, M is called **cographic**. Graphic and cographic matroids are always regular. There exist matroids not linearly representable over a field K , or linearly representable over a particular field K but not regular, or regular but neither graphic nor cographic. If $[G]$ is a matroid isomorphic to

the matroid of the type in example (3), it is called **transversal**. Such a matroid is linearly representable when the cardinality of the field K is sufficiently large.

References

- [1] G. C. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, *Z. Wahrscheinlichkeitstheorie*, 2 (1964), 340–368.
- [2] M. Hall, Jr. Combinatorial theory, Blaisdell, 1967.
- [3] C. L. Liu, Introduction to combinatorial mathematics, McGraw-Hill, 1968.
- [4] H. Whitney, On the abstract properties of linear dependence, *Amer. J. Math.*, 57 (1935), 509–533.
- [5] B. L. van der Waerden, *Algebra*, Ungar, 1970. (Original in German, 1937.)
- [6] D. J. A. Welsh, *Matroid theory*, Academic Press, 1976.
- [7] E. L. Lawler, *Combinatorial optimization—Networks and matroids*, Holt, Rinehart and Winston, 1976.
- [8] M. Iri and S. Fujishige, Use of matroid theory in operations research, circuits and system theory, *Int. J. System Sci.*, 12 (1981), 27–54.
- [9] J. Riordan, *An introduction to combinatorial analysis*, Wiley, 1958.
- [10] H. J. Ryser, *Combinatorial mathematics*, Wiley, 1963.

67 (III.11) Commutative Rings

A. General Remarks

A ring R (\rightarrow 368 Rings) whose multiplication is commutative is called a **commutative ring**. Throughout this article, we mean by a ring a commutative ring with unity element.

B. Ideals

Since our rings are commutative, we need not distinguish right or left ideals from \dagger ideals. A subset \mathfrak{a} of a ring R is an ideal of R if and only if \mathfrak{a} is an R -submodule of R (\rightarrow 277 Modules), if and only if \mathfrak{a} is the \dagger kernel of a ring homomorphism from R into some ring (except for the case $\mathfrak{a} = R$). Given an ideal \mathfrak{a} of a ring R , the set of elements which are \dagger nilpotent modulo \mathfrak{a} , i.e., $\{x \in R \mid x^n \in \mathfrak{a} \ (\exists n)\}$, is called the **radical** of \mathfrak{a} and is often denoted by $\sqrt{\mathfrak{a}}$. The radical of the zero ideal is called the **radical**

of R , or more precisely, the **nilradical** of R (\rightarrow 368 Rings H).

For a subset S of a ring R , the smallest ideal \mathfrak{a} containing S is called the ideal **generated** by S , and S is called a **basis** for \mathfrak{a} ; if S is a finite set, then S is called a **finite basis**. When \mathfrak{a}_λ ($\lambda \in \Lambda$) are ideals of R , the **sum** of these ideals, denoted by $\sum \mathfrak{a}_\lambda$, is defined to be the ideal generated by the union of the ideals \mathfrak{a}_λ ($\lambda \in \Lambda$). If Λ is a finite set, say $\{1, 2, \dots, n\}$, then the sum is denoted also by $\mathfrak{a}_1 + \dots + \mathfrak{a}_n$, and in this case, the sum is also called a finite sum. Note that $\mathfrak{a}_1 + \dots + \mathfrak{a}_n = \{\mathfrak{a}_1 + \dots + \mathfrak{a}_n \mid \mathfrak{a}_i \in \mathfrak{a}_i\}$. The **product** $\mathfrak{a}_1 \dots \mathfrak{a}_n$ of a finite number of ideals $\mathfrak{a}_1, \dots, \mathfrak{a}_n$ is defined to be the ideal generated by the set $\{\mathfrak{a}_1 \dots \mathfrak{a}_n \mid \mathfrak{a}_i \in \mathfrak{a}_i\}$. The intersection of an arbitrary number of ideals is an ideal. When \mathfrak{a} is an ideal of a ring R and S is a subset of R , the **quotient** $\mathfrak{a} : S$ is defined to be the ideal $\{x \in R \mid xS \subset \mathfrak{a}\}$. If $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}_\lambda$ ($\lambda \in \Lambda$) are ideals, we have $(\mathfrak{a} : \mathfrak{b}) : \mathfrak{c} = \mathfrak{a} : \mathfrak{bc}$, $\mathfrak{a} : \sum \mathfrak{d}_\lambda = \bigcap_\lambda (\mathfrak{a} : \mathfrak{d}_\lambda)$.

C. Prime Ideals

An ideal \mathfrak{p} of a ring R is called a **prime ideal** if R/\mathfrak{p} is an \dagger integral domain; \mathfrak{p} is a prime ideal if and only if $\mathfrak{p} \neq R$ and also $ab \in \mathfrak{p}$ ($a, b \in R$) implies $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$ (some literature includes the ring R itself in the set of prime ideals). Let S be a **multiplicatively closed subset** of a ring R , i.e., a nonempty \dagger subsemigroup of R with respect to multiplication. A maximal member among the set of ideals which do not meet S is called a **maximal ideal with respect to S** . Such a member is necessarily a prime ideal; when $S = \{1\}$, it is called a **maximal ideal** of R . An ideal \mathfrak{m} is a maximal ideal of R if and only if R/\mathfrak{m} is a field.

D. Jacobson Radical

The intersection J of all maximal ideals of a ring R is called the **Jacobson radical** of R ; in some cases, this intersection J is called the radical of R (\rightarrow 368 Rings H). Let N be an R -submodule of a finite R -module M . If $MJ + N = M$, then $M = N$ (**Krull-Azumaya lemma** or **Nakayama lemma**).

E. Krull Dimension

For a prime ideal \mathfrak{p} of a ring R , the maximum of the lengths n of \dagger descending chains of prime ideals $\mathfrak{p} = \mathfrak{p}_0 \supsetneq \mathfrak{p}_1 \supsetneq \dots \supsetneq \mathfrak{p}_n$ which begin with \mathfrak{p} (or ∞ if the maximum does not exist) is called the **height** or **rank** of the prime ideal \mathfrak{p} . For an ideal \mathfrak{a} , the minimum of the heights of prime ideals containing \mathfrak{a} is called the **height** of the ideal \mathfrak{a} . The maximum of heights of prime

ideals of R is called the **Krull dimension** (or **altitude**) of the ring R . For an ideal α of R , the Krull dimension of R/α is called the **Krull dimension** (or **depth**) of the ideal α . (The meanings of the terms *rank*, *dimension*, and *depth* now depend on the writings in which they are found; standardization of definition of these terms is becoming more and more of a necessity.)

F. Primary Ideals

For an ideal q of R , if $q \neq R$ and every zero divisor of R/q is nilpotent, then q is called a **primary ideal**. (If R is included in the set of prime ideals, then R is regarded as a primary ideal.) In this case, $p = \sqrt{q}$ is a prime ideal, and q is then said to **belong** to p or to be a **p-primary ideal**. The intersection of a finite number of primary ideals belonging to the same prime ideal p is **p-primary**. Assume that an ideal α is expressed as the intersection $\alpha = q_1 \cap \dots \cap q_n$ of a finite number of primary ideals q_1, \dots, q_n . If this intersection is **irredundant**, that is, if none of the q_i is superfluous in the expression of α , then the set of $\sqrt{q_i}$ ($i = 1, \dots, n$) is uniquely determined by α . The $\sqrt{q_i}$ are called **prime divisors** (or **associated prime ideals**) of the ideal α ; a minimal one among these is called a **minimal** (or **isolated**) **prime divisor** of the ideal α , and those which are not minimal are called **embedded prime divisors** of the ideal α . A maximal one among the prime divisors of α is called a **maximal prime divisor** of the ideal α . (For definitions of these concepts in the case where α is an arbitrary ideal, see [4].) If, in the expression of α as above, n is the smallest occurring in similar expressions, then the expression is called the **shortest representation** of the ideal α by primary ideals. In this case, each q_i is called a **primary component** of the ideal α ; if $\sqrt{q_i}$ is isolated, then q_i is called an **isolated primary component** of the ideal α , and otherwise q_i is called an **embedded primary component** of the ideal α . Isolated primary components are uniquely determined by α , but embedded primary components are not.

G. Rings of Quotients

Let R be a ring. Then the set U of elements of R which are not zero divisors is multiplicatively closed. In the set $R \times U = \{(r, u) | r \in R, u \in U\}$, we define a relation \equiv by $(r, u) \equiv (r', u') \Leftrightarrow ru' = r'u$. Then \equiv is an equivalence relation, and the equivalence class of (r, u) is denoted by r/u . In the set Q of these r/u , addition and multiplication are defined by $r/u + r'/u' = (ru' + r'u)/uu'$, $(r/u)(r'/u') = rr'/uu'$. Then

Q becomes a ring, and $r/1$ can be identified with r . Thus Q is a ring containing R , generated by R and inverses of elements of U . This property characterizes Q , which is called the **ring of total quotients** of the ring R . If R is an integral domain, then Q is a field, called the **field of quotients** of the integral domain R . Let S be a multiplicatively closed subset of R such that $0 \notin S$, and let n be $\{x \in R | xs = 0 (\exists s \in S)\}$ and ϕ be the natural homomorphism $R \rightarrow R/n$. Then none of the elements of $\phi(S)$ is a zero divisor. The subring of the ring of total quotients of R/n , generated by R/n and inverses of elements of $\phi(S)$, is called the **ring of quotients of the ring R with respect to S** , and is denoted by R_S ($R[S^{-1}]$ or RS^{-1}). When M is an R -module, $M \otimes_R R_S$ is called the **module of quotients of the R -module M with respect to S** . It is significant that R_S is tR -flat. There is a one-to-one correspondence between the set of primary ideals q of R which do not meet S and the set of primary ideals \mathfrak{Q} of R_S such that q corresponds to \mathfrak{Q} if and only if $\mathfrak{Q} = qR_S$ ($q = \mathfrak{Q} \cap R$). When p is a prime ideal of R , the complement $R - p$ is multiplicatively closed, and R_{R-p} is called the **local ring of p** or the **ring of quotients of the ring R with respect to the prime ideal p** , and is denoted by R_p (\rightarrow 284 Noetherian Rings C, D). A ring of quotients is also called a **ring of fractions**.

H. Divisibility

In a ring R , if $a = bc$ ($a, b, c \in R$), then we say that b is a **divisor** (or **factor**) of a , and that a is a **multiple** of b , or a is **divisible** by b . We denote this by $b|a$. This relation between $a, b \in R$ is called **divisibility relation** in R . If, in this situation, c has its inverse in R , we say that a is an **associate** of b . A factor b of a is called a **proper factor** if b is neither an associate of a nor invertible. An element which has no proper factor is called an **irreducible element**. A nonzero element which generates a prime ideal is called a **prime element**.

If in an integral domain R every nonzero element is a product of prime elements (up to invertible factors), then we say that the **unique factorization theorem** holds in R , and that R is a **unique factorization domain** (or simply u.f.d.).

Let $A = \{a_1, \dots, a_n\}$ be a set of nonzero elements of a ring R . A **common divisor** of A is an element which is a factor of a_i . A **common multiple** of A is defined similarly. The **greatest common divisor** (G.C.D.) of A is a common divisor which is a multiple of any common divisor; the **least common multiple** (L.C.M.) of A is a common multiple m which is a factor of any common multiple. Thus, the G.C.D. and L.C.M. exist if R is a u.f.d.

I. Integral Dependence

Let R be a subring of a ring R'' sharing a unity element with R'' . An element $a \in R''$ is said to be **integral** (or **integrally dependent**) over R if there are a natural number n and elements c_i of R such that $a^n + c_1 a^{n-1} + \dots + c_n = 0$. If every element of a subset S of R'' is integral over R , we say that S is **integral** over R . (When R has no unity element, a similar definition is given under an additional condition that $c_i \in R^i$. An important special case is where R is an ideal. See D. G. Northcott and D. Rees, *Proc. Cambridge Philos. Soc.*, 50 (1954); M. Nagata, *Mem. Coll. Sci. Univ. Kyôto*, 30 (1956).) The set \tilde{R} of elements of R'' which are integral over R is a ring and is called the **integral closure** of R in R'' . If $\tilde{R} = R$, then R is said to be **integrally closed** in R'' . If R is integrally closed in its ring of total quotients, we say that R is **integrally closed**. An integrally closed integral domain is called a **normal ring**. (In some literature, an integrally closed ring is called a normal ring.) An element $a \in R''$ is called **almost integral** over R if there is an element b of R such that b is not a zero divisor and $a^n b \in R$ for every natural number n . If an element a of the ring of total quotients of R is integral over R , then a is almost integral over R . R is said to be **completely integrally closed** if its ring of total quotients contains no elements which are almost integral over R except the elements of R itself.

J. Group Theorem

Let Q be the ring of total quotients of a ring R . An R -submodule α of Q is called a **fractional ideal** of R if there is a non-zero-divisor c of R such that $c\alpha \subset R$. The product of fractional ideals is defined similarly as in the case of products of ideals. The inverse α^{-1} of the fractional ideal α is defined to be $\{x \in Q \mid x\alpha \subset R\}$. If α contains an element which is not a zero divisor, then α^{-1} is also a fractional ideal. When R is completely integrally closed, we define fractional ideals α and β to be equivalent if $\alpha^{-1} = \beta^{-1}$. This gives rise to an equivalence relation between fractional ideals. The set of equivalence classes of fractional ideals which contain non-zero-divisors forms a group. This result is called the **group theorem**.

An integral domain R is called a **Krull ring** if (i) for every prime ideal \mathfrak{p} of height 1, the ring $R_{\mathfrak{p}}$ is a \dagger discrete valuation ring; (ii) R is the intersection of all the valuation rings $R_{\mathfrak{p}}$ and (iii) every nonzero element a of R is contained in only a finite number of prime ideals of height 1. In a Krull ring R , for an arbitrary nonzero fractional ideal α , there is a uniquely determined product of powers of prime ideals

of height 1 which is equivalent to α in the sense stated above (\rightarrow 439 Valuations).

K. Dedekind Domains and Principal Ideal Domains

A ring R is called a **Dedekind domain** if (i) R is a \dagger Noetherian integral domain, (ii) R is a normal ring, and (iii) the Krull dimension of R is 1. For an integral domain R which is not a field, R is a Dedekind domain if and only if the set of all nonzero fractional ideals is a group, if and only if every nonzero ideal of R is expressed as the product of a finite number of prime ideals and such expression is unique up to the order of prime factors. An important example of a Dedekind domain is the ring of all \dagger algebraic integers, i.e., the \dagger principal order of an \dagger algebraic number field of finite degree. In general, if R is a Dedekind domain with field of quotients K and L is a finite algebraic extension of K , then the integral closure of R in L and any ring R' such that $R \subset R' \subsetneq K$ are Dedekind domains.

An ideal generated by an element is called a **principal ideal**; a fractional ideal generated by an element is called a **principal fractional ideal** (or simply **principal ideal**). In a Dedekind domain, the set P of nonzero principal fractional ideals is a subgroup of the group I of nonzero fractional ideals; I/P is called the **ideal class group** of R , and a member of it is called an **ideal class**. The \dagger order of I/P is called the **class number** of R (\rightarrow 14 Algebraic Number Fields). There are many Dedekind domains whose class numbers are infinite.

A ring R is called a **principal ideal ring** if every ideal is principal; furthermore, if R is an integral domain, then R is called a **principal ideal domain**. A principal ideal ring is the direct sum of a finite number of rings of which each direct summand is either a principal ideal domain or a \dagger local ring whose maximal ideal is a principal nilpotent ideal. A principal ideal domain which is not a field is a Dedekind domain and a u.f.d. We consider, for an arbitrary natural number n and a principal ideal ring R , the set $M(n, R)$ of all $n \times n$ matrices over R . Given an element A of $M(n, R)$, there exist elements X, Y in $M(n, R)$ such that (i) X^{-1}, Y^{-1} are in $M(n, R)$, and (ii) denoting by b_{ij} the (i, j) -entry of XAY , we have $b_{11}R \supset b_{22}R \supset \dots \supset b_{nn}R$ and $b_{ij} = 0$ if $i \neq j$. The nonzero members of the set $\{b_{11}, b_{22}, \dots, b_{nn}\}$ are called the **elementary divisors** of the matrix A . Applying this to a finite module M over the principal ideal ring R , we see that M is the direct sum of $m_1 R, \dots, m_r R$ ($m_i \in M$) such that, with $\alpha_i = \{x \in R \mid m_i x = 0\}$, we have $\alpha_1 \subset \alpha_2 \subset \dots \subset \alpha_r$ (\rightarrow 2 Abelian Groups B; 269 Matrices E).

L. Euclid Rings

A ring R is called a **Euclid ring** if there is a map φ of $R - \{0\}$ into a \dagger well-ordered set W (the set of natural numbers is a special case) satisfying the condition that if $a, b \in R$, $a \neq 0$, then there are $r, q \in R$ such that $b = aq + r$ and either $r = 0$ or $\varphi(r) < \varphi(a)$.

Every Euclid ring is a principal ideal ring. Besides the ring \mathbf{Z} of rational integers, there are several familiar examples of Euclid rings, such as $\mathbf{Z}[\sqrt{-1}]$, $\mathbf{Z}[\sqrt{-2}]$, $\mathbf{Z}[\omega]$ ($\omega^3 = 1$, $\omega \neq 1$), and the polynomial ring of one variable over a field.

References

- [1] W. Krull, *Idealtheorie*, Erg. Math., Springer, second edition, 1968.
- [2] B. L. van der Waerden, *Algebra I, II*, Springer, 1966, 1967.
- [3] O. Zariski and P. Samuel, *Commutative algebra I, II*, Van Nostrand, 1958, 1960.
- [4] M. Nagata, *Local rings*, Interscience, 1962.
- [5] N. Bourbaki, *Eléments de mathématique*, Algèbre, ch. 1, 8, Actualités Sci. Ind., Hermann, 1144b, 1964; 1261a, 1958.
- [6] N. Bourbaki, *Eléments de mathématique*, Algèbre commutative, ch. 1–7, Actualités Sci. Ind., Hermann, ch. 1, 2, 1290a, 1961; ch. 3, 4, 1293a, 1967; ch. 5, 6, 1308, 1964; ch. 7, 1314, 1965.
- [7] A. Grothendieck, *Eléments de géométrie algébrique*, Publ. Math. Inst. HES, 4 (1960).
- [8] D. G. Northcott, *Lessons on rings, modules and multiplicities*, Cambridge Univ. Press, 1968.

68 (XII.10) Compact and Nuclear Operators

A. General Remarks

Let X be a finite-dimensional linear space. Then a linear operator in X is surjective if and only if it is injective. If X is infinite-dimensional, this is no longer the case in general. For an operator of the form $1 + K$ with an integral operator K of continuous kernel, I. Fredholm [1] developed the theory of determinants and retrieved the above equivalence (\rightarrow 217 Integral Equations). Later, F. Riesz [3] simplified the proof and showed that the equivalence holds if K is a compact operator in a Banach space. I. Ts. Gokhberg and M. G. Krein [4] reformulated the result as

the stability of indices of operators. The concept of compact operators itself was introduced by D. Hilbert (who used the terminology, “completely continuous operators”). The Hilbert-Schmidt class is also due to him. These classes of operators are employed in the spectral theory of integral operators [2]. The trace class of operators in Hilbert spaces is defined as the class of operators for which the trace has meaning. A. Grothendieck [5] introduced nuclear operators, extending the definition of trace class operators to general Banach spaces. Nuclear operators and related classes of operators, such as integral operators and absolutely summing operators, play important roles in the theory of topological tensor products [5], determinants [6], vector measures, and measures in linear spaces [7].

B. Compact Operators

A \dagger linear operator T from a \dagger Banach space X to a Banach space Y is said to be **compact** (or **completely continuous**) if T maps any \dagger bounded set of X to a \dagger relatively compact set of Y . In other words, T is compact if, for any bounded sequence $\{x_n\}$ in X , the sequence $\{Tx_n\}$ in Y contains a \dagger strongly convergent subsequence. A compact operator is necessarily \dagger bounded and hence continuous. A compact operator from X to Y maps any \dagger weakly convergent sequence in X to a strongly convergent sequence in Y . If X is \dagger reflexive, the converse is also true.

In this article the set of all bounded (resp. compact) linear operators from X to Y are denoted by $\mathbf{B}(X, Y)$ (resp. $\mathbf{B}^{(c)}(X, Y)$).

C. Examples of Compact Operators

(1) **Degenerate operators.** An operator $T \in \mathbf{B}(X, Y)$ is said to be degenerate or of **finite rank** if the \dagger range $R(T)$ of T is finite-dimensional. A degenerate operator is necessarily compact. The \dagger identity operator in X is compact ($Y = X$) if and only if X is finite-dimensional. (2) **\dagger Integral operators with continuous kernel.** Let E and F be bounded closed regions in the Euclidean spaces \mathbf{R}^m and \mathbf{R}^n , respectively, and let $k = k(t, s)$, $t \in F$, $s \in E$, be a continuous function defined on $F \times E$. Then the integral operator T with \dagger kernel k ,

$$(Tx)(t) = \int_E k(t, s)x(s)ds, \quad t \in F, \quad (1)$$

is a compact operator from the Banach space $C(E)$ to $C(F)$. (For the notation for various function spaces \rightarrow 168 Function Spaces.) (3) **Integral operators of Hilbert-Schmidt type.** In example (2) let E and F be \dagger Lebesgue mea-

surable sets, and let $k \in L_2(F \times E)$. Then the integral operator T determines a compact operator from the Hilbert space $L_2(E)$ to the Hilbert space $L_2(F)$. (4) Let Ω be a bounded open set in \mathbb{R}^n , $s < t$ two real numbers, and $H^s(\Omega)$ the Sobolev space of order s (\rightarrow 168 Function Spaces). Then the natural injection from $H^t(\Omega)$ into $H^s(\Omega)$ is compact (F. Rellich's lemma).

D. Properties of Compact Operators

Any linear combination of compact operators is again compact. If a sequence $\{T_n\}$ of compact operators converges in the norm (i.e., in the uniform operator topology), then the limit T is compact. Thus $\mathbf{B}^{(c)}(X, Y)$ is a closed subspace of the Banach space $\mathbf{B}(X, Y)$ with the operator norm. Any product of a compact operator and a bounded operator is compact. Namely, $A \in \mathbf{B}^{(c)}(X, Y)$, $B \in \mathbf{B}(Y, Z)$, and $C \in \mathbf{B}(Z, X)$ imply $BA \in \mathbf{B}^{(c)}(X, Z)$ and $AC \in \mathbf{B}^{(c)}(Z, Y)$. In particular, $\mathbf{B}^{(c)}(X) = \mathbf{B}^{(c)}(X, X)$ forms a closed two-sided ideal of $\mathbf{B}(X) = \mathbf{B}(X, X)$. An operator $T \in \mathbf{B}(X, Y)$ is compact if and only if its dual operator $T' \in \mathbf{B}(Y', X')$ is compact. The range of a compact operator is always separable. Let X and Y be Hilbert spaces. Then for any $T \in \mathbf{B}^{(c)}(X, Y)$ there exist orthonormal sets $\{\varphi_n\}$ in X and $\{\psi_n\}$ in Y and a sequence $\{c_n\}$ of nonnegative numbers with $\lim c_n = 0$ such that

$$Tu = \sum_n c_n (u, \varphi_n) \psi_n, \quad u \in X.$$

Consequently, any compact operator between Hilbert spaces X, Y can be approximated by a sequence of degenerate operators in the operator norm. However, there are Banach spaces X and Y for which the statement is no longer true. In fact, a Banach space X (resp. the dual Y' of a Banach space Y) has the approximation property (\rightarrow 37 Banach Spaces L) if and only if every $T \in \mathbf{B}^{(c)}(Y, X)$ is the limit in norm of a sequence of degenerate operators for any Banach space Y (resp. X).

E. The Riesz-Schauder Theorem

Let $T \in \mathbf{B}^{(c)}(X)$ and consider a pair of linear equations

$$u - Tu = f, \quad f \in X, \quad (2)$$

$$\varphi - T'\varphi = g, \quad g \in X', \quad (3)$$

where $T' \in \mathbf{B}(X')$ is the dual operator of T in the dual space X' of X . Put $\mathcal{M} = \{u \in X \mid u = Tu\}$ and $\mathcal{M}' = \{\varphi \in X' \mid \varphi = T'\varphi\}$. Then one and only one of the following two cases (i) and (ii) occurs. (i) $\mathcal{M} = \{0\}$, $\mathcal{M}' = \{0\}$; for any $f \in X$

equation (2) has a unique solution; and for any $g \in X'$ equation (3) has a unique solution. (ii) $\dim \mathcal{M} = \dim \mathcal{M}' = m$, $1 \leq m < \infty$; (2) has a solution if and only if f is orthogonal to \mathcal{M}' (i.e., $(f, \varphi) = 0$ for any $\varphi \in \mathcal{M}'$); and (3) has a solution if and only if g is orthogonal to \mathcal{M} (i.e., $(g, u) = 0$ for any $u \in \mathcal{M}$). This is called the **Riesz-Schauder theorem**. In particular, when T is an integral operator in a suitable function space, this theorem is also called **Fredholm's alternative theorem** for integral equation (2).

F. Fredholm Operators and Their Indices

Let T be a closed linear operator from a Banach space X to a Banach space Y . T is called a **Fredholm operator** or an **operator with index** if both $\text{Ker } T$ ($=$ null space $N(T)$) and $\text{Coker } T$ ($= Y/R(T)$) are finite-dimensional. Then the integer $\text{ind } T = \dim \text{Ker } T - \dim \text{Coker } T$ is called the **index** of T . If a closed linear operator T has $\text{Coker } T$ of finite dimension, then the range $R(T)$ is closed. Moreover, if $R(T)$ is closed and the domain $D(T)$ is dense, then $\dim \text{Coker } T = \dim \text{Ker } T'$, where T' is the dual of T , and hence $\text{ind } T = \dim \text{Ker } T - \dim \text{Ker } T'$. A linear operator K from X to Y is said to be **T -compact** if the domain $D(K)$ contains $D(T)$ and if K from $D(T)$ with the graph norm into Y is compact, i.e., for any bounded sequence $\{x_n\}$ in $D(T)$ with $\sup \|Tx_n\| < \infty$, the sequence $\{Kx_n\}$ contains a strongly convergent subsequence.

The following are basic properties of Fredholm operators [4, 11]: (1) If T is Fredholm and K is T -compact, then $T + K$ is Fredholm and $\text{ind}(T + K) = \text{ind } T$. (2) If T is Fredholm and another linear operator S is sufficiently close to T in (graph) norm, then S is also Fredholm and $\text{ind } S = \text{ind } T$. (3) If T is a Fredholm operator from X into Y and S is a Fredholm operator with dense domain in Y into Z , then ST is Fredholm and $\text{ind } ST = \text{ind } S + \text{ind } T$. If $D(T)$ is dense, then $D(ST)$ is also dense. (4) Let T be a closed linear operator with dense domain; then T is Fredholm if and only if the dual T' is Fredholm, and then $\text{ind } T' = -\text{ind } T$.

Fredholm operators and their indices were first studied by Russian mathematicians in connection with boundary value problems of differential equations and singular integral equations. M. F. Atiyah and I. M. Singer have proved that an elliptic linear differential operator P of order m on a compact differentiable manifold M is a Fredholm operator from the Sobolev space $H^s(M)$ into $H^{s-m}(M)$ for any s and that its index is computed from the symbol of P and the characteristic classes of M (\rightarrow 237 K -Theory). Similarly, the indices of linear ordinary differential operators (and

more generally of maximally overdetermined systems of linear partial differential operators) are computed in various spaces of functions and generalized functions (H. Komatsu, B. Malgrange, J.-P. Ramis, M. Kashiwara).

G. Spectra of Compact Operators

The following structure of the \dagger spectrum of compact operators $T \in \mathbf{B}^{(c)}(X)$ is derived from the Riesz-Schauder theorem. The spectrum $\sigma(T)$ of T consists of at most countably many points and has no accumulation points except possibly for 0. Any nonzero point of $\sigma(T)$ is an \dagger eigenvalue of T . When X is infinite-dimensional, 0 always belongs to $\sigma(T)$ but is not necessarily an eigenvalue of T . Each nonzero eigenvalue of T has finite (algebraic) \dagger multiplicity, and hence the \dagger eigenspace $\mathcal{M}_\lambda(T) = \{u \mid Tu = \lambda u\}$, $\lambda \neq 0$, is finite-dimensional. If λ is an eigenvalue of T , then $\bar{\lambda}$ is an eigenvalue of the \dagger adjoint operator T^* of T with the same (either algebraic or geometric) multiplicity as that of λ .

H. Spectral Representations of Compact Normal Operators

Let T be a compact \dagger normal operator in a Hilbert space H . Then we can find a \dagger complete orthonormal set consisting solely of eigenvectors of T . Namely, for each nonzero eigenvalue λ_j of T , take an orthonormal basis $\{\varphi_k^{(j)}\}$ of the eigenspace associated with λ_j . Rearrange all the $\varphi_k^{(j)}$ into a sequence $\{\varphi_n\}$ and add to it, if 0 is an eigenvalue, a complete orthonormal set of the eigenspace associated with 0. Then we obtain a desired complete orthonormal set of H . Let μ_n be the eigenvalue associated with φ_n . Then the sequence $\{\mu_n\}$ is precisely an enumeration of nonzero eigenvalues of T with repetitions according to multiplicity. In terms of $\{\varphi_n\}$ and $\{\mu_n\}$, a \dagger spectral representation of T is given as

$$Tu = \sum_{n=1}^{\infty} \mu_n (u, \varphi_n) \varphi_n, \quad u \in H.$$

The eigenvalue problem of a compact nonnegative \dagger self-adjoint operator T can be solved by means of the following **Rayleigh principle**. Consider **Rayleigh's quotient** $R(x) = (Tx, x) / \|x\|^2$, $x \neq 0$. The largest eigenvalue μ_1 is obtained as $\mu_1 = \max_{x \in X} R(x)$ and any vector x_1 which attains this maximum is an eigenvector. When the largest $n-1$ eigenvalues μ_1, \dots, μ_{n-1} and eigenvectors x_1, \dots, x_{n-1} are determined, the n th eigenvalue μ_n is the maximum of $R(x)$ on the subspace orthogonal to x_1, \dots, x_{n-1} and any vector which attains the

maximum is an eigenvector. A more direct characterization of μ_n , involving no previous eigenvectors, is given by

$$\mu_n = \min_{\substack{f_1, \dots, f_{n-1} \in X \\ f_1, \dots, f_{n-1} \neq 0}} \left(\max_{\substack{(x, f_j) = 0 \\ j=1, \dots, n-1}} R(x) \right).$$

This formula is referred to as the **minimax principle**.

I. Classification of Compact Operators in Hilbert Spaces

Let H and K be Hilbert spaces. Then a $T \in \mathbf{B}(H, K)$ is compact if and only if $(Te_n, f_n) \rightarrow 0$ for any orthonormal sequences e_n in H and f_n in K . J. von Neumann and R. Schatten (*Ann. Math.*, 49 (1948)) classified operators $T \in \mathbf{B}^{(c)}(H, K)$ in the following way. The operator $A = (T^*T)^{1/2}$ is a compact nonnegative self-adjoint operator. Let $\alpha_1 \geq \alpha_2 \geq \dots$ be the enumeration in decreasing order of the positive eigenvalues of A , with each repeated according to its multiplicity. The $\alpha_n = \alpha_n(T)$ are sometimes called the **characteristic numbers** of T . For any $p > 0$ the set of all $T \in \mathbf{B}^{(c)}(H)$ such that

$$\|T\|_p = \left(\sum_{n=1}^{\infty} \alpha_n^p \right)^{1/p} < +\infty$$

is denoted by $\mathbf{B}_p(H)$ (or simply \mathbf{B}_p). Among these classes, \mathbf{B}_1 and \mathbf{B}_2 are most important. \mathbf{B}_1 is called the **trace class** and \mathbf{B}_2 the **Hilbert-Schmidt class**. Correspondingly, $\|T\|_1$ and $\|T\|_2$ are called the **trace norm** and **Hilbert-Schmidt norm** of T , respectively. \mathbf{B}_1 is also called the **nuclear class** and any operator $T \in \mathbf{B}_1$ a **\dagger nuclear operator**. The norm $\|T\|_p$ can also be defined more directly as follows. (i) If $1 < p < \infty$, $\|T\|_p = \sup \| (Te_n, f_n) \|_{l_p}$; (ii) If $0 < p \leq 2$, $\|T\|_p = \inf \| (\|Te_v\|) \|_{l_p}$; (iii) If $2 \leq p < \infty$, $\|T\|_p = \sup \| (\|Te_v\|) \|_{l_p}$, where e_n and f_n range over the orthonormal sequences in H and K , and e_v over the orthonormal bases in H . The class \mathbf{B}_p is a two-sided $*$ ideal in the Banach algebra \mathbf{B} . Moreover precisely, $T \in \mathbf{B}_p$ and $R \in \mathbf{B}$ imply $\|RT\|_p \leq \|R\| \|T\|_p$, $\|TR\|_p \leq \|R\| \|T\|_p$ and $\|T^*\|_p = \|T\|_p$. When $1 \leq p < \infty$, the class \mathbf{B}_p becomes a Banach space with the norm $\|T\|_p$. It is always a \dagger quasi-Banach space, in which the set of all degenerate operators is dense.

The norm $\|T\|_p$ ($T \neq 0$) is a decreasing function of p . Hence $\mathbf{B}_p \subset \mathbf{B}_q$ if $p \leq q$. Also, $T \in \mathbf{B}_p$ and $S \in \mathbf{B}_q$ imply $TS \in \mathbf{B}_r$, where $1/r = 1/p + 1/q$. Let $T \in \mathbf{B}_p$, and let $\{\mu_j\}$ be an enumeration of eigenvalues of T with repetitions according to (either geometric or algebraic) \dagger multiplicity. Then

$$\sum_j |\mu_j|^p \leq \|T\|_p^p.$$

Hilbert-Schmidt class. For an arbitrary complete orthonormal set $\{u_k\}$, we have

$$\|T\|_2^2 = \sum \|Tu_k\|^2.$$

Thus \mathbf{B}_2 can be defined as the set of all $T \in \mathbf{B}$ such that the sum on the right-hand side is finite for a certain complete orthonormal set $\{u_k\}$. A linear operator $T: L_2(E) \rightarrow L_2(F)$ is of Hilbert-Schmidt class if and only if it is an integral operator of Hilbert-Schmidt type (example (3) in Section C). The space \mathbf{B}_2 with the norm $\|T\|_2$ becomes a Hilbert space with the inner product defined by

$$(T, S)_2 = \sum (Tu_k, Su_k),$$

where $\{u_k\}$ is as above.

Trace class. For an operator $T \in \mathbf{B}_1$, the trace $\text{tr}(T)$ of T is defined as

$$\text{tr}(T) = \sum (Tu_k, u_k).$$

Here the right-hand side converges absolutely and does not depend on the complete orthonormal set $\{u_k\}$. The trace is a bounded linear functional on the Banach space \mathbf{B}_1 . The product of two operators of Hilbert-Schmidt type belongs to the trace class, and the converse is also true. If $1 < p, q < \infty$ satisfy $1/p + 1/q = 1$, then the inner product $\langle T, S \rangle$ of $T \in \mathbf{B}_p(H, K)$ and $S \in \mathbf{B}_q(K, H)$ is defined by $\langle T, S \rangle = \text{tr}(ST)$. Under this inner product the dual of the Banach space $\mathbf{B}_p(H, K)$ is identified with $\mathbf{B}_q(K, H)$. Similarly, the dual of $\mathbf{B}^{(c)}(H, K)$ is isomorphic to $\mathbf{B}_1(K, H)$, and the dual of $\mathbf{B}_1(K, H)$ to $\mathbf{B}(H, K)$ (J. Dixmier, Schatten).

J. Volterra Operators

A compact operator T in a Hilbert space is said to be a **Volterra operator** if it is \dagger quasinilpotent, i.e., its \dagger spectral radius is 0. The integral operator

$$(Tx)(t) = \int_t^b k(t, s)x(s)ds, \quad x \in L_2(a, b),$$

appearing in the integral equation of Volterra type is a Volterra operator. Conversely, a Volterra operator satisfying a suitable condition is unitarily equivalent to such an integral operator (M. S. Livshits, Gokhberg and Kreĭn [13]). A Volterra operator admits an abstract triangular representation in a manner similar to Jordan's canonical form (\rightarrow 390 Spectral Analysis of Operators H).

K. Nuclear Operators

Extending the definition of trace class to operators in Banach spaces, A. Grothendieck [5] defined a **nuclear operator** (or **Fredholm**

operator according to [6]) from a Banach space X into a Banach space Y to be a linear operator $T: X \rightarrow Y$ that is represented as $Tx = \sum \lambda_j \langle x, a_j \rangle b_j$ with a sequence $\lambda_j \geq 0$ in l_1 , a bounded sequence a_j in X , and a bounded sequence b_j in Y . In other words, a linear operator $T: X \rightarrow Y$ is nuclear if and only if it is decomposed as the product

$$X \xrightarrow{A} l_\infty \xrightarrow{\Lambda} l_1 \xrightarrow{B} Y, \quad (4)$$

where A and B are bounded linear operators and Λ is multiplication by λ_j in l_1 . The infimum of $\|A\| \|\lambda_j\|_{l_1} \|B\|$ is called the **nuclear norm** of T and is denoted by $\|T\|_1$. When X and Y are Hilbert spaces, this coincides with the trace norm. The integral operator T defined by (1) is nuclear with $\|T\|_1 = \int \sup_t |k(t, s)| ds$. The totality $\mathbf{B}_1(X, Y)$ of nuclear operators $T: X \rightarrow Y$ forms a Banach space under the nuclear norm. If T is nuclear and A is bounded, then $\|TA\|_1 \leq \|T\|_1 \|A\|$ and $\|AT\|_1 \leq \|A\| \|T\|_1$. If T is nuclear, then the dual T' is nuclear, and $\|T'\|_1 \leq \|T\|_1$. Suppose that X and Y are Banach spaces satisfying one of the following conditions: (i) Y is the dual of a Banach space; (ii) X' has the approximation property; (iii) Y'' has the approximation property. Then conversely a $T \in \mathbf{B}(X, Y)$ is nuclear if the dual T' is nuclear, and $\|T\|_1 = \|T'\|_1$. However, this is not necessarily the case in general (T. Figiel and W. B. Johnson).

Replacing l_1 by l_p , we obtain the definition of **operators $T: X \rightarrow Y$ of summable p th power**. Grothendieck [5] considered the case $0 < p \leq 1$ and showed that if T_1, \dots, T_n are of summable p_1, \dots, p_n th power, then the product $T_1 \dots T_n$ is of summable r th power, where r is given by $1/r = (\sum 1/p_i) - (n+1)/2$.

L. Traces and Determinants of Operators

Let T be a linear operator of \dagger finite rank in a linear space X . There are a finite number of elements $a_i \in X'$ and $b_i \in X$ such that

$$Tx = \sum_{i=1}^n \langle x, a_i \rangle b_i. \quad (5)$$

Then $\text{tr}(T) = \sum \langle b_i, a_i \rangle$ and $\det(1 - T) = \det(\delta_{i,j} - \langle b_i, a_j \rangle)$ are independent of the representation (5). Let X be a Banach space. If $T \in \mathbf{B}_1(X)$ is represented as $Tx = \sum \lambda_j \langle x, a_j \rangle b_j$, then it seems reasonable that the **trace** $\text{tr}(T)$ be defined by $\sum \lambda_j \langle b_j, a_j \rangle$, but the sum may depend on the representation. It is known that a Banach space X has the approximation property if and only if the sum does not depend on the representation for any nuclear operator T . However, if T is of summable $(2/3)$ rd power, the trace $\text{tr}(T)$ is always defined uniquely. If X or T satisfies the afore-

mentioned condition, then the **determinant** $\det(1 - T)$ is also defined uniquely as the limit of $\det(1 - T_n)$, where T_n is the n th partial sum. $\det(1 - zT)$ is an entire function of z and its zeros are exactly the reciprocals of the nonzero eigenvalues of T . If $\det(1 - zT) \neq 0$, then the resolvent $(1 - zT)^{-1}$ is given by **Fredholm's formula** extending \dagger Cramer's rule. Eigenvectors are also computed explicitly [6].

Let H be a Hilbert space, and let $T \in \mathbf{B}_1(H)$. Then $\det(1 - zT) = \prod_j (1 - z\lambda_j)$, where λ_j are the nonzero eigenvalues of T , and is an entire function of genus 0. Moreover, let $p \geq 2$ be an integer. Then the modified determinant

$$\det_p(1 - zT) = \prod_j \left[(1 - z\lambda_j) \exp \left(\sum_{k=1}^{p-1} k^{-1} z^k \lambda_j^k \right) \right]$$

is defined for $T \in \mathbf{B}_p(H)$ and is an entire function of genus $p - 1$. This type of determinant was introduced by T. Carleman (*Math. Z.*, 9 (1921)) for the Hilbert-Schmidt class of operators (\rightarrow also Hilbert [2]). It is utilized to prove the completeness of the \dagger root vectors of T and many other facts [9, 12, 13].

M. Weakly Compact Operators

A linear operator from a Banach space X to a Banach space Y is said to be **weakly compact** if T maps any bounded set of X to a relatively weakly compact set. Weakly compact operators are bounded. They have properties similar to those of compact operators. The following are equivalent conditions for a bounded linear operator $T: X \rightarrow Y$. (1) T is weakly compact; (2) $T': Y' \rightarrow X'$ is weakly compact; (3) $T'': X'' \rightarrow Y''$ maps X'' into Y ; (4) There is a reflexive Banach space Z and bounded linear operators $S: X \rightarrow Z$ and $R: Z \rightarrow Y$ such that $T = RS$. The last characterization is due to W. J. Davis, Figiel, Johnson, and A. Pełczyński. Any linear combination of weakly compact operators is weakly compact. The uniform limit of a sequence of weakly compact operators is weakly compact. The product of a weakly compact operator and a bounded operator is weakly compact.

Dunford-Pettis theorem ([9]; R. S. Phillips, *Trans. Amer. Math. Soc.*, 48 (1940)). Let (Ω, μ) be a σ -finite measure space and X a Banach space. If $T: L_1(\Omega) \rightarrow X$ is a weakly compact linear operator, then there exists a bounded strongly measurable function g on Ω with values in X such that

$$Tf = \int f(t)g(t)d\mu(t).$$

Then T maps weakly convergent sequences in $L_1(\Omega)$ into strongly convergent sequences in X . In particular, the product of two weakly com-

pact linear operators in $L_1(\Omega)$ is compact. Similarly, Grothendieck and R. G. Bartle, N. Dunford, and J. Schwartz (\rightarrow [9]) have proved the following: Let Ω be a compact Hausdorff space and X a Banach space. If $T: C(\Omega) \rightarrow X$ is a weakly compact linear mapping, then there is a \dagger vector measure μ defined on the \dagger Borel sets in Ω with values in X such that $\langle \mu, x' \rangle$ is a \dagger Radon measure for any $x' \in X'$ and

$$Tf = \int f(t)d\mu(t).$$

T maps weakly convergent sequences in $C(\Omega)$ into strongly convergent sequences in X . The product of two weakly compact linear operators in $C(\Omega)$ is also compact.

N. Absolutely Summing Operators

A linear operator T from a Banach space X to a Banach space Y is called an **integral operator** if there is a positive \dagger Radon measure μ on the product $A \times B$ of the unit ball A in X' and the unit ball B in Y'' both equipped with the weak* topology such that

$$\langle Tx, y' \rangle = \int \langle x, x' \rangle \langle y', y'' \rangle d\mu(x', y''),$$

or equivalently if $T: X \rightarrow Y''$ is decomposed as (4) with l_∞ and l_1 replaced by $L_\infty(M)$ and $L_1(M)$ for a suitable compact space M with a Radon measure. Nuclear operators are clearly integral. Integral operators are weakly compact, but they are not necessarily compact or nuclear. However, the product of an integral operator and a weakly compact operator is nuclear, and hence every integral operator in a reflexive Banach space is nuclear.

If a compact (resp. weakly compact) operator T from a Banach space X to a Banach space Y maps a closed linear subspace X_1 of X into a closed linear subspace Y_1 of Y , then the restriction $T_1: X_1 \rightarrow Y_1$ and the induced operator $T^1: X/X_1 \rightarrow Y/Y_1$ are also compact (resp. weakly compact). The corresponding result does not hold for nuclear operators and integral operators. The following conditions are equivalent for a bounded linear operator $T: X \rightarrow Y$: (i) There is a Banach space $\tilde{Y} \supset Y$ such that $T: X \rightarrow \tilde{Y}$ is integral; (ii) there is a positive Radon measure μ on the unit ball A of X' such that

$$\|Tx\| \leq \int |\langle x, x' \rangle| d\mu(x');$$

(iii) there is a constant C such that

$$\sum \|Tx_i\| \leq C \sup \{ \sum |\langle x_i, x' \rangle| \mid \|x'\| \leq 1 \}$$

holds for any finite set $\{x_1, \dots, x_n\}$ in X ; (iv) if $\sum x_i$ is an \dagger unconditionally convergent series

in X , then $\sum \|Tx_i\|$ converges. Grothendieck called such an operator a **right semi-integral** in view of property (i), and A. Pietsch called it **absolutely summing** in view of (iv). Similarly, a bounded linear operator $T: X \rightarrow Y$ is called a **left semi-integral** if there is a Banach space \tilde{X} such that $X = \tilde{X}/\tilde{X}_1$ for a closed subspace \tilde{X}_1 , and the product $\tilde{T}: \tilde{X} \rightarrow Y$ is integral. The product of two right (resp. left) semi-integral operators is nuclear. A bounded linear operator in a Hilbert space is right (or left) semi-integral if and only if it is of Hilbert-Schmidt class. A bounded linear operator T from a Banach space X into a Banach space Y is integral if and only if the dual T' is integral. It is a right (resp. left) semi-integral if and only if T' is a left (resp. right) semi-integral.

For other related classes of operators \rightarrow Grothendieck's paper in *Boletim Soc. Mat. São Paulo*, 8 (1953).

References

- [1] I. Fredholm, Sur une classe d'équations fonctionnelles, *Acta Math.*, 27 (1903), 365–390.
- [2] D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Teubner, 1912.
- [3] F. Riesz, Über lineare Funktionalgleichungen, *Acta Math.*, 41 (1918), 71–98.
- [4] I. Ts. Gokhberg (Gohberg) and M. G. Kreĭn, The basic propositions on defect numbers, root numbers and indices of linear operators, *Amer. Math. Soc. Transl.*, ser. 2, 13 (1960), 185–264. (Original in Russian, 1957.)
- [5] A. Grothendieck, Produits tensoriels topologiques et espaces nucléaires, *Mem. Amer. Math. Soc.*, 16 (1955).
- [6] A. Grothendieck, Théorie de Fredholm, *Bull. Soc. Math. France*, 84 (1956), 319–384.
- [7] L. Schwartz, Geometry and probability in Banach spaces, *Lecture notes in math.* 852, Springer, 1980.
- [8] F. Riesz and B. Sz.-Nagy, Leçon d'analyse fonctionnelle, Akad. Kiadó, Budapest, 1952.
- [9] N. Dunford and J. T. Schwartz, *Linear operators I, II*, Interscience, 1958, 1963.
- [10] K. Yosida, *Functional analysis*, Springer, 1965.
- [11] T. Kato, *Perturbation theory for linear operators*, Springer, 1966.
- [12] I. Ts. Gokhberg (Gohberg) and M. G. Kreĭn, Introduction to the theory of linear nonselfadjoint operators, *Amer. Math. Soc.*, 1969. (Original in Russian, 1965.)
- [13] I. Ts. Gokhberg (Gohberg) and M. G. Kreĭn, Theory and applications of Volterra operators in Hilbert space, *Amer. Math. Soc.*, 1970. (Original in Russian, 1967.)

[14] A. Pietsch, *Nuclear locally convex spaces*, Springer, 1972.

69 (IV.8) Compact Groups

A. Compact Groups

A * topological group G is called a **compact group** if the underlying topological space of G is a * compact Hausdorff space. The * torus group $\mathbf{T}^n = \mathbf{R}^n/\mathbf{Z}^n$ ($n = 1, 2, \dots$) (commutative group), the * orthogonal group $O(n)$, the * unitary group $U(n)$, the * symplectic group $Sp(n)$, and the additive group \mathbf{Z}_p of * p -adic integers are compact groups (\rightarrow 60 Classical Groups; for other compact Lie groups \rightarrow 249 Lie Groups, 248 Lie Algebras). Let $C(G)$ be the * linear space formed by all the complex-valued continuous functions f, g, h, \dots defined on a compact group G ; $C(G)$ is a * Banach space with the norm

$$\|f\| = \sup_{x \in G} |f(x)|.$$

Since a compact group G is * locally compact, there exists a right-invariant * Haar measure on G . Because of the compactness of G , the total measure of G is finite, and the measure is also left-invariant. By the condition that the total measure is 1, such a measure is uniquely determined. The integral of f in $C(G)$ relative to this measure is called the **mean value** of f . Since for $f, g \in C(G)$, $f(xy^{-1})g(y)$ is continuous in two variables x, y , the * convolution $f \times g(x) = \int f(xy^{-1})g(y)dy$ also belongs to $C(G)$. $C(G)$ constitutes a ring under the multiplication defined by the convolution. This ring can be considered as an extension of the notion of a * group ring for finite groups; it is called the **group ring** of the compact group G . The function $x \rightarrow f(x^{-1})$ will be denoted by f^* , and the inner product in the * function space $L_2(G)$ will be written as (f, g) .

B. Representations of Compact Groups

Let $G(E)$ be the group of units of * bounded linear operators on a Banach space E , and suppose that we have a homomorphism U of a topological group G into $G(E)$. The homomorphism U is called a **strongly (weakly) continuous representation** on E of G if, for any $a \in E$, the map $x \rightarrow U(x)a$ of G into E is continuous with respect to the * strong (* weak) topology on E . When E is a * Hilbert space, a strongly continuous representation U such that every

$U(a)$ is a unitary operator is called a \dagger unitary representation. Let U be any strongly continuous representation of a compact group G on a Hilbert space E ; for a, b in E , let $\langle a, b \rangle$ be the mean value of the \dagger inner product $(U(x)a, U(x)b)$ on E . Then U is a unitary representation on the Hilbert space E with the new inner product $\langle a, b \rangle$.

A representation U of G on a Banach space E is said to be **irreducible** if E contains no closed subspace other than $\{0\}$ and E , which is invariant under every $U(x)$ ($x \in G$). If a weakly continuous representation of a compact group on a Banach space E is irreducible, then E is finite-dimensional. Moreover, any unitary representation U of a compact group on a Hilbert space E can be decomposed into a discrete \dagger direct sum of irreducible representations. Namely, there exists a family $\{E_\alpha\}_{\alpha \in A}$ of irreducible (hence finite-dimensional) invariant subspaces E_α of E which are orthogonal to each other such that $E = \sum_{\alpha \in A} E_\alpha$. In particular, any continuous representation of a compact group on a finite-dimensional space is \dagger completely reducible. In $L_2(G)$ with respect to the \dagger Haar measure on a \dagger locally compact group G , the representation U defined by $(U(x)f)(y) = f(yx)$ ($x, y \in G, f \in L_2(G)$) is a unitary representation of G . This representation U is called the (right) **regular representation** of G . Decomposition of the regular representation of a compact group G into irreducible representations is given by the **Peter-Weyl theory**, described later.

In the rest of this article, the representations under consideration will be (continuous) representations by matrices of finite degree. Let $D_1(x) = (d_{ij}^{(1)}(x))$ and $D_2(x) = (d_{ij}^{(2)}(x))$ be irreducible unitary representations which are not mutually \dagger equivalent. Then from \dagger Schur's lemma follow the orthogonality relations $(d_{ij}^{(1)}, d_{kl}^{(2)}) = 0$ and $(\sqrt{n_1} d_{ij}^{(1)}, \sqrt{n_1} d_{mn}^{(1)}) = \delta_{im} \delta_{jn}$ (where n_1 is the degree of the representation D_1). From each \dagger class D_α of irreducible representations of G , choose a unitary representative $D_\alpha(x) = (d_{ij}^\alpha(x))$, and let n_α denote its degree. Then from the orthogonality relations it follows that the $\sqrt{n_\alpha} d_{ij}^\alpha(x)$ form an \dagger orthonormal system of $L_2(G)$.

Let $h \in C(G)$ and consider the map $H: f \rightarrow h \times f$ of $C(G)$ to $C(G)$. Then H is a \dagger compact operator in $C(G)$. Since $C(G)$ is contained in $L_2(G)$, we can define the inner product in $C(G)$. By $(h \times f, g) = (f, h^* \times g)$, $h = h^*$ implies $(Hf, g) = (f, Hg)$; that is, H is a \dagger Hermitian operator. For a given f in $C(G)$, there exists $h (= h^*)$ in $C(G)$ such that $h \times f$ is uniformly arbitrarily near f . From the theory of compact Hermitian operators, Hf can be uniformly approximated by linear combinations of the \dagger eigenfunctions of H . Since the \dagger eigenspace of H is finite-

dimensional and invariant under $U(a)$, the eigenfunctions of H are linear combinations of a finite number of $\sqrt{n_\alpha} d_{ij}^\alpha(x)$. Hence any function f in $C(G)$ can be uniformly approximated by linear combinations of a finite number of $\sqrt{n_\alpha} d_{ij}^\alpha(x)$. This fact, like the similar result in \dagger Fourier series, is called the **approximation theorem**. From this, it follows that the orthonormal system $\{\sqrt{n_\alpha} d_{ij}^\alpha(x)\}$ is \dagger complete; i.e., if an element of $C(G)$ is orthogonal to each element in this system, then it is 0.

Since $C(G)$ is dense in $L_2(G)$, $\{\sqrt{n_\alpha} d_{ij}^\alpha(x)\}$ is a \dagger complete orthonormal system of the Hilbert space $L_2(G)$. Hence for any $f \in L_2(G)$, its "Fourier series" $\sum_\alpha \sum_{ij} c_{ij}^\alpha \sqrt{n_\alpha} d_{ij}^\alpha(x)$ (where $c_{ij}^\alpha = (f, \sqrt{n_\alpha} d_{ij}^\alpha)$) converges to f in the mean of order 2 (i.e., with respect to the \dagger metric of $L_2(G)$). In particular, if G is a compact \dagger Lie group and f is sufficiently many times differentiable, then this series converges uniformly to f .

The space V_i^α of dimension n_α spanned by the elements $d_{ij}^\alpha(x)$ ($1 \leq j \leq n_\alpha$) in the i th row of the matrix $D_\alpha(x)$ is invariant under the right regular representation U . The representation on V_i^α given by U can only be D_α . Then the fact that $\{\sqrt{n_\alpha} d_{ij}^\alpha(x)\}$ is an orthonormal system of $L_2(G)$ means that the regular representation U of a compact group G is decomposable into a discrete direct sum of finite-dimensional irreducible representations. Each irreducible representation D_α is contained in U with multiplicity equal to its degree n_α .

If a function $\varphi(x)$ in $C(G)$ satisfies $\varphi(y^{-1}xy) = \varphi(x)$ for any x, y , then it is called a **class function**. The set $K(G)$ of all the continuous class functions coincides with the \dagger center of the group ring $C(G)$. The \dagger character of an irreducible representation of G is a class function, and the set $\{\chi_\alpha(x)\}$ of all characters plays the same role as the orthogonal system $\{\sqrt{n_\alpha} d_{ij}^\alpha(x)\}$ in $C(G)$. Namely, $\{\chi_\alpha(x)\}$ is a complete orthonormal system in (the \dagger completion of) $K(G)$, and any class function can be uniformly approximated by linear combinations of a finite number of these characters.

The preceding paragraphs give a brief description of the Peter-Weyl theory. If G is the 1-dimensional \dagger torus group $\mathbf{T}^1 = \mathbf{R}/\mathbf{Z}$, namely, the compact group of real numbers mod 1, then this is actually the theory of Fourier series concerning periodic functions on the line. (For concrete irreducible representations of $O(n)$, $U(n)$, $Sp(n)$, and formulas for characters \rightarrow 60 Classical Groups. For representations of compact Lie groups \rightarrow 249 Lie Groups, 248 Lie Algebras.) The theory of compact groups was completed by F. Peter and H. Weyl (*Math. Ann.*, 97 (1927)), and J. von Neumann's theory concerning almost periodic functions in a group (1934) united the

theory of compact groups with H. Bohr's theory of almost periodic functions (\rightarrow 18 Almost Periodic Functions).

C. Structure of Compact Groups

Let a be an element of a compact group G different from e . Since the underlying space of a topological group is a \dagger completely regular space, there exists a function in $C(G)$ such that $f(a) \neq f(e)$. Hence there exists a representation $D(x)$ of G such that $D(a)$ is not equal to the unit matrix. This means that any compact group G can be expressed as a \dagger projective limit group of compact Lie groups. Beginning with this fact, von Neumann (1933) showed that a \dagger locally Euclidean compact group is a Lie group (\rightarrow 423 Topological Groups N).

D. Set of Representations

The set $G' = \{D\}$ of representations of G by matrices admits the following operations: (i) $D_1 \otimes D_2$ (\dagger tensor product representation); (ii) $D_1 \oplus D_2 = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}$ (\dagger direct sum representation); (iii) $P^{-1}DP$ (equivalent representation); and (iv) \bar{D} (complex conjugate representation). Let M be a subset of G' such that $\bar{D} \in M$ whenever $D \in M$ and the irreducible components of $D_1 \otimes D_2$ are in M whenever $D_1, D_2 \in M$. Then M is called a **module of representations** of G . There is a one-to-one correspondence between closed normal subgroups H of G and modules M formed by all the representations of G/H .

A representation of G' is a correspondence which assigns to each D a matrix $A(D)$ of the same degree as that of D and preserves the operations of G' : $A(D_1 \otimes D_2) = A(D_1) \otimes A(D_2)$, $A(D_1 \oplus D_2) = A(D_1) \oplus A(D_2)$, $A(P^{-1}DP) = P^{-1}A(D)P$, and $A(\bar{D}) = \overline{A(D)}$. Let G'' be the set of all the representations of G' . Define the product of $A_1, A_2 \in G''$ by $A_1 A_2(D) = A_1(D)A_2(D)$ and a topology on G'' by the \dagger weak topology of the functions $A(D)$ of D . Namely, a typical neighborhood of A_0 is of the form $U(A_0; D_1, \dots, D_s; \varepsilon) = \{A \mid \|A(D_i) - A_0(D_i)\| < \varepsilon, i = 1, \dots, s\}$. G'' is a topological group under this multiplication and topology. Then the **Tannaka duality theorem** states that $G'' \cong G$ holds (T. Tannaka, *Tôhoku Math. J.*, 45 (1939)). Let $R(G)$ be the \dagger algebra over the complex number field \mathbb{C} formed by the set of all the linear combinations of a finite number of $d_{ij}^{\lambda}(x)$, and let $\text{Aut } R(G)$ be the automorphism group of this algebra $R(G)$. Let G^* be the set of all the elements σ in $\text{Aut } R(G)$ which commute with every left translation $L(x)$ ($(L(x)f)(y) = f(xy)$) and which satisfy

$\sigma(\bar{f}) = \overline{\sigma(f)}$. Then G^* is a topological group with respect to the weak topology, and the Tannaka duality theorem implies that the correspondence which assigns to each $x \in G$ the right translation $R(x)$ (the restriction of $U(x)$ to $R(G)$) is an isomorphism of G onto G^* as topological groups. For the case where G is a compact Lie group, C. Chevalley restated the Tannaka duality theorem as one giving a relation between compact Lie groups and complex algebraic groups (\rightarrow 249 Lie Groups U).

References

See references to 18 Almost Periodic Functions and 423 Topological Groups.
For compact Lie groups \rightarrow references to 249 Lie Groups.

70 (IX.5) Complexes

A. General Remarks

The notion of complexes was introduced by H. Poincaré to study the topology of \dagger manifolds by combinatorial methods (\rightarrow 65 Combinatorial Manifolds). Various kinds of complexes have been introduced in the course of the development of topology. These will be dealt with individually in the sections that follow.

B. Euclidean Simplicial Complexes

Let \mathbb{R}^N be the N -dimensional \dagger Euclidean space, and let a_0, a_1, \dots, a_m be points of \mathbb{R}^N with coordinates $a_i = (a_i^{(1)}, a_i^{(2)}, \dots, a_i^{(N)})$ for $i = 0, 1, \dots, m$. For real numbers $\lambda_0, \lambda_1, \dots, \lambda_m$, we denote by $\lambda_0 a_0 + \lambda_1 a_1 + \dots + \lambda_m a_m$ the point $(\sum_{i=0}^m \lambda_i a_i^{(1)}, \sum_{i=0}^m \lambda_i a_i^{(2)}, \dots, \sum_{i=0}^m \lambda_i a_i^{(N)})$ of \mathbb{R}^N .

A set of $n+1$ points a_0, a_1, \dots, a_n of \mathbb{R}^N is said to be **independent** or in **general position** if vectors $\overrightarrow{a_0 a_1}, \overrightarrow{a_0 a_2}, \dots, \overrightarrow{a_0 a_n}$ are linearly independent. For given independent points a_0, a_1, \dots, a_n , let $|a_0 a_1 \dots a_n|$ denote the subset of \mathbb{R}^N given by

$$|a_0 a_1 \dots a_n| = \{\lambda_0 a_0 + \lambda_1 a_1 + \dots + \lambda_n a_n \mid \lambda_0 + \lambda_1 + \dots + \lambda_n = 1, \lambda_i \geq 0\}.$$

$|a_0 a_1 \dots a_n|$ is called an n -**simplex** with **vertices** a_0, a_1, \dots, a_n . Simplexes are denoted by $\Delta, \Delta', \sigma, \tau$, etc. For example, a 0-simplex $|a_0|$ is a point a_0 , a 1-simplex $|a_0 a_1|$ is a segment $\overline{a_0 a_1}$, and a 2-simplex $|a_0 a_1 a_2|$ is a triangle $\Delta a_0 a_1 a_2$.

We attach to $\lambda_0 a_0 + \lambda_1 a_1 + \dots + \lambda_n a_n$ the coordinates $(\lambda_0, \lambda_1, \dots, \lambda_n)$, called the **barycentric coordinates**. If $|a_0 a_1 \dots a_n|$ is an n -simplex, every subset $\{a_{i_0}, a_{i_1}, \dots, a_{i_q}\}$ of $\{a_0, a_1, \dots, a_n\}$ is independent. A q -simplex $|a_{i_0} a_{i_1} \dots a_{i_q}|$ is called a q -**face** of $|a_0 a_1 \dots a_n|$.

A set \mathfrak{R} of simplexes in a Euclidean space \mathbf{R}^N is called a **Euclidean (simplicial) complex** in \mathbf{R}^N if \mathfrak{R} satisfies the following three conditions: (i) Every face of a simplex belonging to \mathfrak{R} is also an element of \mathfrak{R} . (ii) The intersection of two simplexes belonging to \mathfrak{R} is either empty or a face of each of them. (iii) Each point of a simplex belonging to \mathfrak{R} has a neighborhood in \mathbf{R}^N that intersects only a finite number of simplexes belonging to \mathfrak{R} . A Euclidean complex is also called a **geometric complex** (or **rectilinear complex**). Each 0-simplex in \mathfrak{R} is called a **vertex** in \mathfrak{R} . We define the **dimension** of \mathfrak{R} to be n if \mathfrak{R} contains an n -simplex but no $(n+1)$ -simplex, and ∞ if \mathfrak{R} contains n -simplexes for all $n \geq 0$.

By a **subcomplex** of a Euclidean complex \mathfrak{R} we mean a Euclidean complex that is a subset of \mathfrak{R} . For a Euclidean complex \mathfrak{R} , its r -**section** (or r -**skeleton**) is defined to be the subcomplex of \mathfrak{R} consisting of all n -simplexes ($n \leq r$) in \mathfrak{R} .

If \mathfrak{R} is a Euclidean complex in \mathbf{R}^N , we denote by $|\mathfrak{R}|$ the set of points in \mathbf{R}^N belonging to simplexes in \mathfrak{R} . This set $|\mathfrak{R}|$ is called the **Euclidean polyhedron** of \mathfrak{R} .

By **subdivision** \mathfrak{R}' of a Euclidean complex \mathfrak{R} we mean a Euclidean complex such that $|\mathfrak{R}'| = |\mathfrak{R}|$ and each simplex in \mathfrak{R}' is contained in a simplex in \mathfrak{R} .

Specifically, we can construct a subdivision \mathfrak{R}' of \mathfrak{R} utilizing \dagger barycenters of simplexes in \mathfrak{R} ; namely, we let \mathfrak{R}' be the set of all r -simplexes whose vertices consist of barycenters of the series $\Delta_0 \subset \Delta_1 \subset \dots \subset \Delta_r$ of simplexes in \mathfrak{R} . Then \mathfrak{R}' is a subdivision of \mathfrak{R} , called the **barycentric subdivision** of \mathfrak{R} and denoted by $\text{Sd } \mathfrak{R}$.

Given a Euclidean complex \mathfrak{R} and a subset A of $|\mathfrak{R}|$, we define the **star** of A in \mathfrak{R} to be the subcomplex of \mathfrak{R} that consists of simplexes $\{\Delta\}$ and their faces such that $\Delta \cap A \neq \emptyset$. Furthermore, we define the **open star** of A in \mathfrak{R} as the union of \dagger open simplexes (the interiors of simplexes) of \mathfrak{R} whose closures intersect A . We denote by $\text{St}_{\mathfrak{R}}(A)$ the star of A in \mathfrak{R} and by $O_{\mathfrak{R}}(A)$ the open star of A in \mathfrak{R} ; then $O_{\mathfrak{R}}(A)$ is an open set whose closure is $|\text{St}_{\mathfrak{R}}(A)|$.

The notion of Euclidean simplicial complexes can be generalized to that of **Euclidean cell complexes**; this is done by replacing the term *simplex* by \dagger *convex cell* in the definition of Euclidean simplicial complex. For a Euclidean cell complex \mathfrak{R} , the notions of vertex, dimension, subcomplex, r -section, and subdivision are defined similarly as in the case of Euclidean simplicial complexes.

C. Simplicial Complexes

Given a Euclidean simplicial complex \mathfrak{R} , let K denote the set of all the vertices in \mathfrak{R} , and let Σ denote the set consisting of those subsets $\{v_0, \dots, v_r\}$ of K for which there exist simplexes Δ in \mathfrak{R} such that $\{v_0, \dots, v_r\}$ coincide with the set of vertices of Δ . Then we have (1) if $s \in \Sigma$ and $s \supset s'$, $s' \neq \emptyset$, then $s' \in \Sigma$; (2) every set consisting of a single element in K is in Σ , and the empty set is not in Σ .

In general, if a pair (K, Σ) of a set K and a set Σ consisting of finite subsets of K satisfy (1) and (2), then the pair (or the set K) is called an **abstract simplicial complex** (or simply **simplicial complex**). If K is a simplicial complex, each element of the set K is called a **vertex** in K , and each set of Σ is called a **simplex** in K . A simplex consisting of $n+1$ vertices is called an n -**simplex**. We say that a simplicial complex K is **finite** if it consists of a finite number of vertices; it is **locally finite** if every vertex of K belongs to only finitely many simplexes in K . We define similarly **countable simplicial complexes** and **locally countable simplicial complexes**. The **dimension** and r -**section** of a simplicial complex are defined as in the case of Euclidean complexes. By a **subcomplex** of a simplicial complex K we mean a simplicial complex K_0 such that each simplex of K_0 is a simplex in K .

Let K and L be simplicial complexes. A mapping $\varphi: K \rightarrow L$ is called a **simplicial mapping** (**simplicial map**) of K to L if the following condition is satisfied: If v_0, v_1, \dots, v_n are vertices of a simplex of K , then $\varphi(v_0), \varphi(v_1), \dots, \varphi(v_n)$ are vertices of some simplex of L . Two simplicial complexes K and L are said to be **isomorphic** if there exist simplicial mappings $\varphi: K \rightarrow L, \psi: L \rightarrow K$ such that $\psi \circ \varphi$ and $\varphi \circ \psi$ are the identity mappings.

Given a simplicial complex K , let $|K|$ denote the set of all functions x from the set of vertices of K to the closed interval $I = [0, 1]$ satisfying the following conditions: (i) The set $\{v \in K | x(v) \neq 0\}$ is a simplex of K . (ii) $\sum_{v \in K} x(v) = 1$. The value $x(v)$ is called the **barycentric coordinate** of the point $x \in |K|$ with respect to the vertex v . Each vertex v of K is identified with the point of $|K|$ whose barycentric coordinate with respect to the vertex v is 1 and is called a **vertex** in $|K|$. For a simplex $s = \{v_0, v_1, \dots, v_n\}$ in K , we define $|s| = \{x \in |K| | x(v) = 0 (v \notin s)\}$, which is called a **simplex** in $|K|$. We call $\{x \in |s| | x(v_i) > 0 (i = 0, 1, \dots, n)\}$ an **open simplex** of $|K|$ or the **interior** of $|s|$. We remark here that for an arbitrary simplex s in K a point $x \in |s|$ can be written in the form $x = \sum_{v \in s} x(v) \cdot v$. We can define a metric d on $|K|$ by $d(x, y) = (\sum_{v \in K} (x(v) - y(v))^2)^{1/2}$. However, $|K|$ is usually supplied with a \dagger stronger topology

defined as follows: (1) Each simplex $|s|$ in $|K|$ has the topology given by the metric d . (2) A subset U of $|K|$ is open if and only if $U \cap |s|$ is an open subset of $|s|$ for each simplex s of K . Henceforward, by the topology of $|K|$ we mean the topology just defined unless otherwise stated. The set $|K|$ with such a topology is called the **polyhedron** of K . The topology of $|K|$ coincides with the above metric topology if and only if K is locally finite. If simplicial complexes K and L are isomorphic, then $|K|$ and $|L|$ are homeomorphic. If K is the simplicial complex defined by a Euclidean simplicial complex \mathfrak{K} , then $|K|$ and $|\mathfrak{K}|$ are homeomorphic. When K is finite, there exists an Euclidean simplicial complex \mathfrak{K} whose simplicial complex is isomorphic to K . Accordingly $|K|$ is homeomorphic to the Euclidean polyhedron $|\mathfrak{K}|$.

If K and L are simplicial complexes, a mapping $f: |K| \rightarrow |L|$ satisfying the following condition is said to be **linear**: If $s = \{v_0, v_1, \dots, v_n\}$ is a simplex in K and $x = \lambda_0 v_0 + \dots + \lambda_n v_n$ ($\lambda_0 + \dots + \lambda_n = 1, \lambda_i \geq 0$), then $f(v_0), \dots, f(v_n)$ belong to a simplex in L and $f(x) = \lambda_0 f(v_0) + \dots + \lambda_n f(v_n)$. The linear mapping determined by a simplicial mapping $\varphi: K \rightarrow L$ is denoted by $|\varphi|: |K| \rightarrow |L|$ and is also called a **simplicial mapping** and denoted by the same letter φ . Let K and K' be simplicial complexes. If there exists a linear mapping $l: |K'| \rightarrow |K|$, which is a homeomorphism, then we identify $|K|$ and $|K'|$ by l and call K' a **subdivision** of K . The **barycentric subdivision** $Sd K$ of a simplicial complex K is defined as in the case of Euclidean complexes. We also have notions of **star** $St_K(A)$ and **open star** $O_K(A)$ for a simplicial complex K and a subset A of $|K|$. If K and L are simplicial complexes, a mapping $f: |K| \rightarrow |L|$ is called a **piecewise linear mapping** if there exist subdivisions K' and L' of K and L , respectively, such that $f: |K'| \rightarrow |L'|$ is linear.

Given an open \dagger covering $\mathfrak{M} = \{M_v\}_{v \in K}$ of a topological space X , the index set K becomes a simplicial complex if we consider each finite nonempty subset s of K such that $\bigcap_{v \in s} M_v \neq \emptyset$ (empty) to be a simplex. The resulting simplicial complex K is called the **nerve** of the open covering \mathfrak{M} . Furthermore, if $\mathfrak{M} = \{M_v\}_{v \in K}$ and $\mathfrak{N} = \{N_w\}_{w \in L}$ are open coverings of a set X , \mathfrak{N} is a refinement of the covering \mathfrak{M} , and L is the nerve of \mathfrak{N} , then a simplicial mapping $\varphi: L \rightarrow K$ is defined by sending each vertex w in L to a vertex v in K such that $N_w \subset M_v$.

Given two disjoint simplicial complexes K and L , a simplicial complex $K * L$, called the **join** of K and L , is defined by the following: (1) the vertices of $K * L$ are the vertices of K and the vertices of L . (2) A nonempty subset of vertices is a simplex of $K * L$ if and only if its

subsets in K and L are empty or simplexes there. In particular, the join of a simplicial complex K and a single point is called the **cone** of K .

A simplicial complex K is said to be **ordered** if a \dagger partial ordering is given in the set of vertices in K such that the set of vertices of each simplex is \dagger totally ordered. Given ordered simplicial complexes K and L , an ordered simplicial complex $K \times L$, called the **Cartesian product** of K and L , is defined by the following:

(1) The vertices in $K \times L$ are pairs (v, w) , where v and w are any vertices in K and L , respectively. (2) A set of vertices $(v_0, w_0), \dots, (v_n, w_n)$ such that $v_0 \leq \dots \leq v_n$ and $w_0 \leq \dots \leq w_n$ is a simplex in $K \times L$ if (v_0, \dots, v_n) and (w_0, \dots, w_n) are simplexes in K and L , respectively; all simplexes in $K \times L$ are obtained in this manner. (3) $(v_1, w_1) \leq (v_2, w_2)$ if and only if $v_1 \leq v_2$ and $w_1 \leq w_2$.

Assume that either K or L is locally finite or that both K and L are locally countable. Let X and Y be topological spaces and I be the closed interval $[0, 1]$. We define an equivalence relation \sim in the topological space $X \cup (X \times Y \times I) \cup Y$ by $x \sim (x, y, 0)$ and $y \sim (x, y, 1)$, where $x \in X, y \in Y$. The quotient space of $X \cup (X \times Y \times I) \cup Y$ by this relation is called the **join** of X and Y and is denoted by $X * Y$. Then the polyhedron $|K \times L|$ is homeomorphic to the product space $|K| \times |L|$ of the topological spaces $|K|$ and $|L|$. And the polyhedron $|K * L|$ is homeomorphic to $|K| * |L|$.

By a **triangulation** T of a topological space X we mean a pair (K, t) consisting of a simplicial complex K and a homeomorphism $t: |K| \rightarrow X$. A triangulation is also called a **simplicial decomposition**. If $T = (K, t)$ is a triangulation of X , the various concepts defined for K can be transferred to X by means of the mapping t . For example, by a **simplex** of the triangulation T we mean the image of a simplex of $|K|$ under t . We say that a triangulation $T = (K, t)$ is **finite** if K is a finite simplicial complex. If $T = (K, t)$ is a triangulation and (K', l) is a subdivision of K , then $T' = (K', t \circ l)$ is called a **subdivision** of T . If $T_1 = (K_1, t_1), T_2 = (K_2, t_2)$ are triangulations of topological spaces X_1, X_2 , respectively, a mapping $f: X_1 \rightarrow X_2$ is called a **simplicial mapping** relative to T_1 and T_2 if $t_2^{-1} \circ f \circ t_1: |K_1| \rightarrow |K_2|$ is a simplicial mapping. The following two problems on triangulations are famous: (1) Under what topological conditions is it possible for a given topological space to be supplied with a triangulation? (2) Given two triangulations T_1, T_2 of a space X , are there subdivisions $T'_1 = (K'_1, t'_1), T'_2 = (K'_2, t'_2)$ of T_1 and T_2 , respectively, such that K'_1 and K'_2 are isomorphic? Concerning the second problem, the conjecture asserting the existence of subdivisions T'_1

and T'_2 as above is known as the **fundamental conjecture (Hauptvermutung)** in topology. Every 3-dimensional manifold is triangulable, and any two of its triangulations admit subdivisions satisfying the condition in (2) (E. E. Moise [7]).

In 1961, J. Milnor showed that the fundamental conjecture is not true for polyhedra [8]. The triangulation problem and the fundamental conjecture for topological manifolds were negatively solved by R. Kirby and L. Siebenmann [9] (\rightarrow 65 Combinatorial Manifolds). Any \dagger differentiable manifold is triangulable, and the fundamental conjecture holds for its $\dagger C^r$ -triangulations ($r \geq 1$) [6] (\rightarrow 114 Differential Topology).

Let T_1, T_2 be triangulations of topological spaces X_1, X_2 , respectively, and let $f: X_1 \rightarrow X_2$ be a continuous mapping. Then a simplicial mapping $\varphi: X_1 \rightarrow X_2$ relative to T_1 and T_2 is called a **simplicial approximation** to f if, for each $x \in X$, the image $\varphi(x)$ lies on the simplex of T_2 whose interior contains $f(x)$. The following existence theorem is called the **simplicial approximation theorem**: For every continuous mapping $f: X_1 \rightarrow X_2$, there exist a subdivision T'_1 of T_1 and a simplicial mapping $\varphi: X_1 \rightarrow X_2$ relative to T'_1 and T_2 that is a simplicial approximation of f . If the triangulation T_1 is finite, then for a sufficiently large n we can choose $\text{Sd}^n T_1$ as the T'_1 above (where $\text{Sd}^0 T = T$ and $\text{Sd}^n T = \text{Sd}(\text{Sd}^{n-1} T)$ ($n \geq 1$)). If $\varphi: X_1 \rightarrow X_2$ is a simplicial approximation to a continuous mapping f , then f and φ are \dagger homotopic.

D. Cell Complexes

Let V^n be the \dagger unit n -disk, S^{n-1} be the \dagger unit $(n-1)$ -sphere, and X be a Hausdorff space. For a subset e of X , let \bar{e} be the closure of e in X , and let $\dot{e} = \bar{e} - e$. A subset e of the space X is called an n -cell, or **open n -cell** in X if there is a relative homeomorphism $\varphi: (V^n, S^{n-1}) \rightarrow (\bar{e}, \dot{e})$, i.e., a continuous mapping $\varphi: V^n \rightarrow \bar{e}$ such that $\varphi(S^{n-1}) \subset \dot{e}$ and $\varphi: V^n - S^{n-1} \rightarrow \bar{e} - \dot{e}$ is a homeomorphism. For example, $S^n - \{p\}$ ($p \in S^n$) is an n -cell. A set $\{e_\lambda | \lambda \in \Lambda\}$ of cells in the Hausdorff space X is called a **cellular decomposition** of X if the following three conditions are satisfied: (i) $e_\lambda \cap e_\mu$ is empty if $\lambda \neq \mu$; (ii) $X = \bigcup_{\lambda \in \Lambda} e_\lambda$; (iii) If the dimension of e_λ is $n+1$, then $\dot{e}_\lambda \subset X^n$, where X^n is the union of all the cells e_μ ($\mu \in \Lambda$) whose dimensions are not greater than n . For example, the n -sphere S^n has a cellular decomposition consisting of a single 0-cell and a single n -cell.

A Hausdorff space X together with its cellular decomposition $\{e_\lambda\}$ is called a **cell complex**, and each e_λ is called a **cell** in the cell

complex X . For a cell complex, the notions of **vertex**, **n -section**, and **dimension** are defined in the same way as the corresponding notions in Euclidean complexes. Let X be a cell complex and A a topological subspace of X such that the closure of each cell of X intersecting A is contained in A . Then the set of cells e such that $e \cap A \neq \emptyset$ forms a cellular decomposition of A . The set A together with this cellular decomposition is called a **subcomplex** of the cell complex X . A cell complex X with its cells $\{e_\lambda\}$ is said to be **finite** if the number of e_λ is finite. If each point in a cell complex X is an interior point of some finite subcomplex of X , then X is said to be **locally finite**. We define similarly a **countable cell complex** and a **locally countable cell complex**. If the closure of each n -cell of a cell complex X is homeomorphic to V^n ($n=0, 1, \dots$), X is said to be **regular**. If X and Y are cell complexes, a continuous mapping $f: X \rightarrow Y$ such that $f(X^n) \subset Y^n$ ($n=0, 1, \dots$) is called a **cellular mapping**. If X and Y are cell complexes, the set of \dagger topological products $e_1 \times e_2$, where e_1, e_2 run over all cells of X, Y , respectively, is a cellular decomposition of the product space $X \times Y$. The resulting cell complex $X \times Y$ is called the **product complex** of the cell complexes X and Y .

A cell complex X is said to be **closure finite** if each cell in X is contained in a finite subcomplex of X ; and X is said to have the **weak topology** if a subset $U \subset X$ is open if and only if $U \cap \bar{e}$ is relatively open in \bar{e} for each cell e of X . We call a cell complex a **CW complex** if it is closure finite and has the weak topology. The cellular decomposition of a CW complex is called a **CW decomposition**. A locally finite cell complex is a CW complex.

Fundamental properties of CW complexes are as follows: (i) A CW complex is a \dagger paracompact (hence \dagger normal) space and is \dagger locally contractible. (ii) A subcomplex A of a CW complex X is a closed subspace of X , and A itself is a CW complex. (iii) A mapping $f: X \rightarrow Y$ of a CW complex X to a topological space Y is continuous if and only if the restriction $f|_{\bar{e}}$ is continuous for each cell e of X . (iv) If X and Y are CW complexes and $f: X \rightarrow Y$ is any continuous mapping then there exists a cellular mapping of X to Y that is homotopic to f (**cellular approximation theorem**). (v) A pair (X, A) consisting of a CW complex X and its subcomplex A has the \dagger homotopy extension property for any topological space. (vi) A CW complex has the \dagger covering homotopy property for any \dagger fiber space. (vii) The product complex $X \times Y$ of two CW complexes X and Y is not necessarily a CW complex, but it is \dagger homotopy equivalent to a CW complex. (viii) If either X or Y is locally finite, or if both X and Y are locally countable, then the product

complex $X \times Y$ is a CW complex. (ix) For CW complexes X and Y , the mapping space Y^X is homotopy equivalent to a CW complex. (x) The covering space of a CW complex has a CW decomposition.

If K is a simplicial complex, the polyhedron $|K|$ is a regular CW complex whose cells are all open simplexes in $|K|$. A simplicial complex K is (locally) finite if and only if the CW complex $|K|$ is (locally) finite. In particular, the Euclidean polyhedron of a Euclidean simplicial (or cell) complex is a locally finite CW complex. A polyhedron $|K|$ generally admits a CW decomposition whose cells are far smaller in number than the simplexes constituting a simplicial decomposition of K . For any CW complex X , there exists a polyhedron $|K|$ that is homotopy equivalent to X . In particular, if X is an n -dimensional finite (countable) CW complex, we can choose as K an n -dimensional finite (locally finite and countable) simplicial complex.

E. Semisimplicial Complexes

By an **ordered simplex** in a simplicial complex K we mean a finite sequence (v_0, v_1, \dots, v_n) ($n \geq 0$) of vertices in K , contained in the set of vertices of a simplex in K . Let $O(K)_n$ be the set of all ordered simplexes of K of length $n+1$, and define mappings $\partial_i: O(K)_n \rightarrow O(K)_{n-1}$ and $s_i: O(K)_n \rightarrow O(K)_{n+1}$ for $i=0, 1, \dots, n$ by $\partial_i(v_0, \dots, v_n) = (v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$ and $s_i(v_0, \dots, v_n) = (v_0, \dots, v_{i-1}, v_i, v_i, v_{i+1}, \dots, v_n)$. Then the following relations hold:

$$\begin{aligned} \partial_i \circ \partial_j &= \partial_{j-1} \circ \partial_i \quad (i < j), \\ s_i \circ s_j &= s_{j+1} \circ s_i \quad (i \leq j), \\ \partial_i \circ s_j &= s_{j-1} \circ \partial_i \quad (i < j), \\ \partial_i \circ s_j &= s_j \circ \partial_{i-1} \quad (i > j+1), \\ \partial_i \circ s_i &= \partial_{i+1} \circ s_i = \text{identity}. \end{aligned} \quad (1)$$

Let Δ^n be the n -dimensional simplex in \mathbf{R}^n with vertices $e_0 = (0, 0, \dots, 0)$, $e_1 = (1, 0, \dots, 0)$, \dots , $e_n = (0, \dots, 1)$. By a **singular n -simplex** in a topological space X we mean a continuous mapping $T: \Delta^n \rightarrow X$. Let $S(X)_n$ be the set of all singular n -simplexes in X , and define mappings $\partial_i: S(X)_n \rightarrow S(X)_{n-1}$ and $s_i: S(X)_n \rightarrow S(X)_{n+1}$ for $i=0, 1, \dots, n$ by $\partial_i T(\lambda_0, \dots, \lambda_n) = T(\lambda_0, \dots, \lambda_{i-1}, 0, \lambda_i, \dots, \lambda_n)$ and $s_i T(\lambda_0, \dots, \lambda_{n+1}) = T(\lambda_0, \dots, \lambda_{i-1}, \lambda_i + \lambda_{i+1}, \lambda_{i+1}, \dots, \lambda_{n+1})$, where $(\lambda_0, \dots, \lambda_n)$ is the point $\sum_{i=0}^n \lambda_i e_i$, $\lambda_i \geq 0$, $\sum_{i=0}^n \lambda_i = 1$. Then relation (1) holds between ∂_i and s_i .

Because of the importance of relation (1), which is basic in defining homology of simplicial complexes and topological spaces (\rightarrow 201 Homology Theory), S. Eilenberg and J. A. Zilber gave the following definition: A **semi-**

simplicial complex K consists of a sequence of sets K_n ($n=0, 1, \dots$) together with mappings $\partial_i: K_n \rightarrow K_{n-1}$, $s_i: K_n \rightarrow K_{n+1}$ ($i=0, 1, \dots, n$) satisfying relations (1). An element of K_n is called an **n -simplex** in K , and ∂_i , s_i are called the **i th face operator** and the **i th degeneracy operator**, respectively. A simplex is said to be **degenerate** if it is the image of a simplex under some s_i . A semisimplicial complex is abbreviated as **s.s. complex**. The s.s. complexes $O(K) = \{O(K)_n, \partial_i, s_i\}$ and $S(X) = \{S(X)_n, \partial_i, s_i\}$ are called the **ordered complex** of K and the **singular complex** of X , respectively.

Let K be an s.s. complex, and let L_n be a subset of K_n for $n=0, 1, \dots$. If $\partial_i(L_n) \subset L_{n-1}$ and $s_i(L_n) \subset L_{n+1}$ for each i , then $L = \{L_n, \partial_i|L_n, s_i|L_n\}$ is an s.s. complex, and L is called a **subcomplex** of the s.s. complex K . If A is a subspace of a topological space X , $S(A)$ is a subcomplex of $S(X)$. If K is an ordered simplicial complex, a subcomplex $O'(K)$ of $O(K)$ is obtained by considering the set of all ordered simplexes (v_0, v_1, \dots, v_n) such that $v_0 \leq v_1 \leq \dots \leq v_n$ ($n=0, 1, \dots$). If K and L are s.s. complexes, a sequence $f = \{f_n\}$ of mappings $f_n: K_n \rightarrow L_n$ defined for each n is called an **s.s. mapping** if $\partial_i \circ f_n = f_{n-1} \circ \partial_i$ and $s_i \circ f_n = f_{n+1} \circ s_i$ ($0 \leq i \leq n$). If $f: X \rightarrow Y$ is a continuous mapping of topological spaces, then f determines an s.s. mapping $S(f): S(X) \rightarrow S(Y)$ by $S(f)(T) = f \circ T$. Two s.s. complexes K and L are said to be **isomorphic** if there is a bijective s.s. mapping of K to L . For two s.s. complexes K and L , we define the **Cartesian product** $K \times L$ to be the s.s. complex given by $(K \times L)_n = K_n \times L_n$, $\partial_i(\sigma, \tau) = (\partial_i \sigma, \partial_i \tau)$, $s_i(\sigma, \tau) = (s_i \sigma, s_i \tau)$ ($\sigma \in K_n, \tau \in L_n$). If K and L are ordered simplicial complexes, the s.s. complexes $O'(K) \times O'(L)$ and $O'(K \times L)$ are isomorphic. If X and Y are topological spaces, the s.s. complexes $S(X) \times S(Y)$ and $S(X \times Y)$ are isomorphic.

Given an s.s. complex K , we construct a topological space $|K|$ as follows: First, we provide K_n with the discrete topology and consider the topological space $\bar{K} = \bigcup_{n \geq 0} K_n \times \Delta^n$. Next we consider simplicial mappings $\varepsilon^i: \Delta^{n-1} \rightarrow \Delta^n$ and $\eta^i: \Delta^{n+1} \rightarrow \Delta^n$ defined by $\varepsilon^i(p_j) = p_j$ ($j < i$), $\varepsilon^i(p_j) = p_{j+1}$ ($j \geq i$) and $\eta^i(p_j) = p_j$ ($j \leq i$), $\eta^i(p_j) = p_{j-1}$ ($j > i$), where p_0, \dots, p_n are the vertices of Δ^n . The topological space $|K|$ is defined to be the quotient space of \bar{K} with respect to an equivalence relation \sim that is defined by the following: $(\partial_i \sigma, y) \sim (\sigma, \varepsilon^i(y))$ ($\sigma \in K_n, y \in \Delta^{n-1}$), $(s_i \sigma, y) \sim (\sigma, \eta^i(y))$ ($\sigma \in K_n, y \in \Delta^{n+1}$), where $i=0, 1, \dots, n$. The space $|K|$ is called the **(geometric) realization of the s.s. complex** K . Given an s.s. mapping $f: K \rightarrow L$, we obtain a continuous mapping $|f|: |K| \rightarrow |L|$ defined by $|f|((\sigma, y)) = |f(\sigma), y|$, where $|\sigma, y|$ is the point in $|K|$ represented by $(\sigma, y) \in \bar{K}$. We call $|f|$ the **realization of the s.s. mapping** f .

The realization $|K|$ of an s.s. complex K is a CW complex whose cells are in one-to-one correspondence with the nondegenerate simplexes in K . For a topological space X , a \dagger weak homotopy equivalence $\rho: |S(X)| \rightarrow X$ is defined by $\rho(|T, y|) = T(y)$ ($T \in S(X)_n, y \in \Delta^n$). This mapping gives rise to a homotopy equivalence when X is a CW complex.

The singular complex $S(X)$ of a topological space X has the following property: Given simplexes $\sigma_0, \dots, \sigma_{k-1}, \sigma_{k+1}, \dots, \sigma_{n+1} \in K_n$ with $\partial_i \sigma_j = \partial_{j-1} \sigma_i$ ($i < j, i, j \neq k$), there exists a simplex $\sigma \in K_{n+1}$ with $\partial_i \sigma = \sigma_i$ ($i \neq k$). An s.s. complex K with this property is called a **Kan complex**. If, in addition, $\partial_i \sigma = \partial_i \sigma'$ ($\sigma, \sigma' \in K_n, i \neq k$) imply $\partial_k \sigma = \partial_k \sigma'$, we call K a **minimal complex**. For every Kan complex K , there are minimal subcomplexes M of K that are isomorphic to each other. Moreover, $|M|$ is a \dagger deformation retract of $|K|$. For a Kan complex K , the \dagger homotopy group can be defined combinatorially.

Two CW complexes X and Y are homotopy equivalent if and only if the minimal subcomplexes of $S(X)$ and $S(Y)$ are isomorphic.

F. Eilenberg-MacLane Complex

Given an integer $n \geq 1$ and a group π (Abelian if $n \geq 2$), there exists an \dagger arcwise connected topological space X for which the \dagger homotopy groups $\pi_i(X)$ are trivial for $i \neq n$ and $\pi_n(X) \cong \pi$. Such a space is called an **Eilenberg-MacLane space** of type (π, n) . Let $\Omega(X; X, *)$ be the \dagger path space over X , and let $p_0: \Omega(X; X, *) \rightarrow X$ be the natural projection. If X is an Eilenberg-MacLane space of type (π, n) , then $(\Omega(X; X, *), p_0, X)$ gives rise to a standard \dagger contractible \dagger fiber space whose fiber is an Eilenberg-MacLane space of type $(\pi, n-1)$. Assume that X is an Eilenberg MacLane space of type (π, n) with Abelian group π . Then X is $(n-1)$ -connected; hence the \dagger Hurewicz theorem can be utilized to show the existence of an isomorphism $h: \pi_n(X) \cong H_n(X)$, while the universal coefficient theorem can be utilized to show that $H^n(X; \pi) \cong \text{Hom}(H_n(X), \pi)$. Since in this case we have $\pi_n(X) \cong \pi$, the element $h^{-1} \in \text{Hom}(H_n(X), \pi_n(X))$ can be regarded as an element of $\text{Hom}(H_n(X), \pi)$. Now the **fundamental class** of X is defined to be the cohomology class $u \in H^n(X; \pi)$ corresponding to h^{-1} . Let Y be a \dagger CW complex and $\pi(Y; X)$ be the set of \dagger homotopy classes of continuous mappings from Y to X . Then there exists a one-to-one correspondence $\pi(Y; X) \rightarrow H^n(Y; \pi)$ given by the assignment $[f] \rightarrow f^* u$ (\rightarrow 305 Obstructions). Let $S(X)$ be the singular complex of X , and let $M(X)$ be a minimal complex of $S(X)$. If X is an Eilenberg-MacLane space

of type (π, n) , $M(X)$ is isomorphic to a certain complex determined uniquely by π and n . This complex is called the **Eilenberg-MacLane complex** of type (π, n) and is denoted by $K(\pi, n)$. The notation $K(\pi, n)$ is also used to mean the space X itself.

Let $\Delta(q)$ be a simplicial complex whose simplexes are all nonvacuous subsets of $\{0, 1, \dots, q\}$. Let $\varepsilon_i: \Delta(q-1) \rightarrow \Delta(q)$ be the simplicial mapping defined by $\varepsilon_i(j) = j$ ($0 \leq j \leq i-1$), $\varepsilon_i(j) = j+1$ ($i \leq j \leq q-1$), and let $\eta_i: \Delta(q+1) \rightarrow \Delta(q)$ be the mapping defined by $\eta_i(j) = j$ ($0 \leq j \leq i$), $\eta_i(j) = j-1$ ($i+1 \leq j \leq q+1$). Now $K(\pi, n)$ is a Kan complex defined by $K(\pi, n)_q = Z^n(\Delta(q); \pi)$, $\partial_i \sigma = \sigma \circ \varepsilon_i$, $s_i \sigma = \sigma \circ \eta_i$, where $Z^0(\Delta(q); \pi) = \pi$, $Z^1(\Delta(q); \pi)$ is the set of π -valued functions defined on the set of pairs (i, j) such that $0 \leq i < j \leq q$ and satisfying the equality $\sigma(j, k) \cdot \sigma(i, k)^{-1} \cdot \sigma(i, j) = 1$ for $0 \leq i < j < k \leq q$, and $Z^n(\Delta(q); \pi)$ ($n \geq 2$) is the group of \dagger oriented cocycles of the simplicial complex $\Delta(q)$. If π is Abelian, the structure of the Abelian group $Z^n(\Delta(q); \pi)$ gives $K(\pi, n)$ the structure of an Abelian group in the s.s. category. This structure yields a one-to-one correspondence $K(\pi, n)_q \rightarrow K(\pi, n-1)_{q-1} \times \dots \times K(\pi, n-1)_0$ for $n \geq 1$ and leads to the expression of $\tau \in K(\pi, n)_q$ in the form $\langle \sigma_{q-1}, \dots, \sigma_0 \rangle$ with $\sigma_i \in K(\pi, n-1)_i$. The **W-construction** of $K(\pi, n-1)$ for $n \geq 1$ is a Kan complex $W(\pi, n-1)$ defined by $W(\pi, n-1)_q = K(\pi, n-1)_q \times K(\pi, n)_q$ and $\partial_0(\sigma_q \times \tau_q) = (\partial_0 \sigma_q) \cdot \sigma_{q-1} \times \partial_0 \tau_q$, $\partial_i(\sigma_q \times \tau_q) = \partial_i \sigma_q \times \partial_i \tau_q$ for $1 \leq i \leq q$, $s_i(\sigma_q \times \tau_q) = s_i \sigma_q \times s_i \tau_q$, where $\sigma_q \in K(\pi, n-1)_q$ and $\tau_q = \langle \sigma_{q-1}, \dots, \sigma_0 \rangle \in K(\pi, n)_q$. Let $p: W(\pi, n-1) \rightarrow K(\pi, n)$ be a natural projection. Then $(W(\pi, n-1), p, K(\pi, n))$ plays the role of the \dagger universal bundle for $K(\pi, n-1)$ in the s.s. category in the following sense: Let L be an s.s. complex, and let $f: L \rightarrow K(\pi, n)$ be an s.s. mapping. We define $f^* W(\pi, n-1)$ to be the subcomplex of $W(\pi, n-1) \times L$ generated by simplexes $(\sigma_q \times \tau_q) \times \rho_q$ such that $\tau_q = f(\rho_q)$, where $\sigma_q \times \tau_q \in W(\pi, n-1)_q$ and $\rho_q \in L_q$. Let $p: f^* W(\pi, n-1) \rightarrow L$ be the natural projection. Then $(f^* W(\pi, n-1), p, L)$ is called the principal fiber bundle induced from $W(\pi, n-1)$ by f . Any principal bundle over L with group $K(\pi, n-1)$ can be expressed as an induced bundle. This property means that $(W(\pi, n-1), p, K(\pi, n))$ is universal. On the other hand, we have an algebraic analog of the universal bundle for the chain group of $K(\pi, n-1)$, called the **bar construction** [12]. Both these concepts were defined by Eilenberg and MacLane in order to determine the structure of the (co)homology of $K(\pi, n)$, which is denoted by $(H^*(\pi, n))H_*(\pi, n)$. This object was later achieved by H. Cartan, who introduced an improved notion called Cartan construction (\rightarrow Appendix A, Table 6.III). Let $\pi(L, K(\pi, n))$ be the set of s.s. homotopy classes of s.s. map-

pings from L to $K(\pi, n)$. If π is Abelian, there exists a one-to-one correspondence $\pi(L, K(\pi, n)) \rightarrow H^n(L; \pi)$ given by the assignment $[f] \rightarrow f^*u$, where $u \in H^n(\pi, n; \pi)$ is the fundamental class of $K(\pi, n)$. By virtue of this correspondence, $k = f^*u \in H^n(L; \pi)$ determines an induced bundle $f^*W(\pi, n-1)$ uniquely up to equivalence, which is denoted by $K(\pi, n-1) \times_k L$.

Let X be an arcwise connected topological space. Let X^n be the Cartesian product of n copies of X . Clearly, the symmetric group \mathfrak{S}_n of degree n operates on X^n . The n -fold **symmetric product** $SP^n X$ of X is defined to be the quotient space of X^n under the action of \mathfrak{S}_n . If we specify a reference point of X , we have a natural inclusion $SP^{n-1} X \subset SP^n X$ and can consider the inductive limit space $\bigcup_{1 \leq n} SP^n X$, denoted by $SP^\infty X$. Then the Dold-Thom theorem [15] shows that $M(SP^\infty X) \cong \prod_{i=1}^\infty K(H_i(X), i)$. In particular, we have $M(SP^\infty S^n) \cong K(\mathbb{Z}, n)$ for $n \geq 1$. This result can be applied to obtain a direct relationship between the axiomatic definition of t cohomology operations using $K(\pi, n)$ due to Eilenberg and Serre and the constructive definition using the symmetric groups due to Steenrod (A. Dold [16], T. Nakamura [17]). For a detailed study of the (co)homology of $SP^n X \rightarrow [18]$.

G. Postnikov Complex

Let X be an arcwise connected topological space. For the sake of simplicity, throughout this section we assume that X is t simple (\rightarrow 202 Homotopy Theory). Then the t Postnikov system of X can be defined as an inverse system (X_n, p_n) ($n=0, 1, 2, \dots$) consisting of topological spaces X_n , continuous mappings $p_n: X_n \rightarrow X_{n-1}$, and a system (X, q_n) ($n=0, 1, 2, \dots$) consisting of continuous mappings $q_n: X \rightarrow X_n$ such that $p_n \circ q_n = q_{n-1}$ and satisfying the following three properties: (1) X_0 is one point. (2) (X_n, p_n, X_{n-1}) is a t fiber space induced from a standard contractible fiber space over an Eilenberg-MacLane space of type $(\pi_n(X), n+1)$ by a mapping corresponding to a cohomology class $k^{n+1} \in H^{n+1}(X_{n-1}; \pi_n(X))$. (3) $q_{n*}: \pi_i(X) \rightarrow \pi_i(X_n)$ gives an isomorphism for $0 \leq i \leq n$. These cohomology classes k^{n+1} are called **Eilenberg-Postnikov invariants** (or simply **k-invariants**). Corresponding to the above facts, the minimal complex $M(X)$ can be obtained as the inverse limit of a certain inverse system $(K(n), p(n))$ consisting of Kan complexes $K(n)$ and s.s. mappings $p(n): K(n) \rightarrow K(n-1)$ defined by $K(0) = K(0, 0)$ and $K(n) = K(\pi_n, n) \times_{k^{n+1}} K(n-1)$ for $n \geq 1$ with natural

projections $p(n)$, where $\pi_n = \pi_n(X)$ and $k^{n+1} \in H^{n+1}(K(n-1); \pi_n)$. This system is determined uniquely up to s.s. homotopy equivalence by its limit, called the **Postnikov complex** and denoted by $K(\pi_1, k^3, \pi_2, \dots, k^{n+1}, \pi_n, \dots)$. As yet we are ignorant of an effective method of computing the cohomology of a Postnikov complex from π_n and k^{n+1} .

References

- [1] H. Cartan, Sur la théorie de Kan, Sém. H. Cartan, Ecole Norm. Sup., 1956–1957.
- [2] S. Eilenberg and N. E. Steenrod, Foundations of algebraic topology, Princeton Univ. Press, 1952.
- [3] P. J. Hilton and S. Wylie, Homology theory, Cambridge Univ. Press, 1960.
- [4] E. H. Spanier, Algebraic topology, McGraw-Hill, 1966.
- [5] R. E. Mosher and M. C. Tangora, Cohomology operations and applications in homotopy theory, Harper & Row, 1968.
- [6] J. R. Munkres, Elementary differential topology, Princeton Univ. Press, second edition, 1966.
- [7] E. E. Moise, Geometric topology in dimensions 2 and 3, Springer, 1977.
- [8] J. Milnor, Two complexes that are homeomorphic but combinatorially distinct, Ann. Math., (2) 74 (1961), 575–590.
- [9] R. C. Kirby and L. C. Siebenmann, Foundational essays on topological manifolds, smoothings and triangulations, Princeton Univ. Press, 1977.
- [10] J. P. May, Simplicial objects in algebraic topology, Van Nostrand, 1967.
- [11] E. Curtis, Simplicial homotopy theory, Lecture notes, Aarhus Univ., 1967.
- [12] S. Eilenberg and S. MacLane, On the groups $H(\pi, n)$ I, II, III, Ann. Math., (2) 58 (1953), 55–106; 60 (1954), 49–139; 60 (1954), 513–557.
- [13] J.-P. Serre, Cohomologie modulo 2 des complexes d'Eilenberg-MacLane, Comment. Math. Helv., 27 (1953), 198–232.
- [14] Sém. H. Cartan, Ecole Norm. Sup., 1954–1955.
- [15] A. Dold and R. Thom, Quasifaserungen und unendliche symmetrische Produkte, Ann. Math., (2) 67 (1958), 239–281.
- [16] A. Dold, Über die Steenrodschen Kohomologieoperationen, Ann. Math., (2) 73 (1961), 258–294.
- [17] T. Nakamura, On cohomology operations, Japan. J. Math., 33 (1963), 93–145.
- [18] M. Nakaoka, Decomposition theorem for homology groups of symmetric groups, Ann. Math., (2) 71 (1960), 16–42.

[19] G. W. Whitehead, Elements of homotopy theory, Springer, 1978.

71 (XVI.9) Complexity of Computations

A. Measures for Complexity of Computation

Intuitively, **complexity of computation** means the amount of computing efforts measured on some suitable scale. When a problem can be solved by means of any one of several algorithms, it is highly desirable to compare the complexities of those algorithms. For example, the power x^{10} can be evaluated with a pocket calculator by using either of the algorithms

$$x^{10} = x \times x \times x \times x \times x \times x \times x \times x \times x \times x, \quad (A)$$

$$u = x \times x, \quad v = u \times u, \quad w = v \times v,$$

$$x^{10} = u \times w. \quad (B)$$

The complexities of these algorithms can be measured by various quantities: (*T1*) time in seconds, (*T2*) the number of times keys and buttons are touched, (*T3*) the number of basic operations (here, multiplications), and (*S*) the number of values to be stored in the calculator. The quantities (*T1*), (*T2*), and (*T3*) are called **time complexities**, and (*S*) is called the **space complexity**. Obviously, algorithm (B) is preferable with respect to time complexity, although (A) is better with respect to space complexity. (If the calculator has only one memory register, then only (A) can be executed without recording numbers by hand.) For many calculators, the following algorithm is best with respect to both time and space complexities:

$$x^{10} = ((x^2)^2 \times x)^2. \quad (C)$$

In the general theory of complexity, the number of basic operations (*T3*) is often taken as the basis of time complexity, since it represents the intrinsic complexity of the algorithm, rather than of extrinsic factors such as human skill or mechanical performance.

For the evaluation of a power x^n in general, the complexity depends on the value of n : such a parameter dominating the complexity is called the **size** of the problem. Let $T_X(n)$ be the number of multiplications required for evaluating x^n by an algorithm X . Let \mathcal{F} be the set of all algorithms for evaluating x^n . The time complexity $T(n)$ for evaluation of the power is then defined by

$$T(n) = \min_{X \in \mathcal{F}} T_X(n).$$

This is satisfied when [6]

$$\log_2 n \leq T(n) \leq 2 \log_2 n$$

and

$$\lim_{n \rightarrow \infty} \frac{T(n)}{\log_2 n} = 1.$$

In a problem concerning a ^{*}graph, the number m of nodes in the graph, the number n of its edges, or their sum $m+n$ are chosen as the size of the graph. However, the complexity is not determined by such a size, because there are many different instances (particular graphs) of the same size. So the complexity of this problem is defined in the following manner. Let $T(I)$ be the complexity of solving an instance I of the problem and $L(n)$ be the set of all possible instances of size n . Then

$$W(n) = \max_{I \in L(n)} T(I)$$

and

$$A(n) = \sum p(I) T(I),$$

where $p(I)$ denotes the relative frequency of the instance I . The value $W(n)$ is called the **worst-case complexity**, and $A(n)$ is called the **average complexity** of the problem.

In some cases the complexity of a circuit is measured by the number of its building blocks [1], and the complexity of a program is measured by its length [2, 3].

B. Complexity of a Decision Problem

A decision problem and its complexity are rigorously formulated in terms of ^{*}Turing machines. Let Σ be a finite set of symbols and Σ^+ be the whole set of nonempty strings of symbols in Σ . Every instance of the problem is assumed to be represented by a string in Σ^+ . A **decision problem** is a triple (Σ, L, P) of the set Σ of symbols, a subset L of Σ^+ , and a mapping P from L to $\{0 \text{ (false), } 1 \text{ (true)}\}$. The set L represents the set of all possible instances of the problem. The problem is said to be **solvable** if and only if there exists a Turing machine M satisfying the following condition for every string α in the set L : When a tape containing the string α is given to the machine M whose head is initially put on the leftmost symbol of α , the machine writes the value of $P(\alpha)$ on the tape and stops after a finite number of steps. Such a Turing machine is said to **compute** the function P . The number $T_M(\alpha)$ of steps is the time complexity of solving the instance α by the machine M , and the length $S_M(\alpha)$ of the used area of the tape is the space complexity. The worst-case complexity $T_M(n)$ is defined by

$$T_M(n) = \max_{\alpha \in L, |\alpha|=n} T_M(\alpha).$$

The space complexity $S_M(n)$ for the size n is defined in a similar way. The complexities $T(n)$ and $S(n)$ of the problem are not defined, because improvement by a constant factor is always possible for any Turing machine [5].

A problem is said to be solved in **linear** or **polynomial time** if its time complexity $T_M(n)$ is bounded by a linear or polynomial function of the size n .

C. Objectives of Complexity Theory

The objectives of complexity theory are as follows:

- (i) Analyze an algorithm X for a problem P and evaluate its complexity. When exact evaluation is hard, the order of magnitude $O(T_X(n))$ is investigated.
- (ii) Construct better algorithms for the problem P . This gives a better upper bound to the complexity $T(n)$ of the problem.
- (iii) Make clear the limitation of improving algorithms for the problem P . This is done by establishing a lower bound to the complexity $T(n)$ of the problem.

A common approach to constructing better algorithms is to partition the problem into smaller parts, apply algorithms to the parts, and then combine the solutions for the parts into a solution for the whole (divide and conquer, [5]). This approach often yields an efficient algorithm for the problem, especially when the partitioning can be repeated recursively. For demonstrating the limitations of algorithms of a certain class, diagonalization [5], determination of information-theoretic lower bounds [6], or the oracle method [6] are often utilized.

D. Elementary Results on Time Complexity

(1) **Number of arithmetic operations.** A polynomial of degree n with one variable can be evaluated in about $3n/2$ arithmetic operations if preconditioning on coefficients is allowed. Otherwise, $2n-1$ operations are necessary and sufficient for evaluating a polynomial of degree n . For computing the product of two square matrices of degree n , $O(n^2)$ operations are necessary and $O(n^{2.5})$ operations are sufficient. The inverse matrix and the value of the determinant of a square matrix of degree n are computed with the same order of operations as the product. Discrete Fourier transformation of n points is executed in $O(n \log n)$ arithmetic operations (fast Fourier transformation).

(2) **Number of comparisons and data transfers.** Rearrangement of n items in increasing order is realized in $O(n \log n)$ comparisons and

data transfers. Selection of the k th largest item among n items is done in $O(n)$ operations. The position of an item in a list containing n items is found in $O(n)$ comparisons by linear search, in $O(\log n)$ comparisons by binary chopping, and in $O(1)$ comparisons on average by the hashing method.

Many other results can be found in [5]–[7].

E. NP-Completeness

A problem is easily solved if its complexity is $O(n)$ or less. The problem is very hard if its complexity is $O(2^n)$, unless the size n is small. Between these types of problems there is a class of problems, each of which can be solved in polynomial time by a Turing machine.

A **nondeterministic Turing machine** is said to **solve** a decision problem (Σ, L, P) if it can detect the case when $P(\alpha) = 1$ by exercising good choices; more precisely, if it can stop after a finite number of steps and write the value 1 of $P(\alpha)$ on the tape, starting from an initial state with the head on the leftmost symbol of a string α on the tape, provided that $P(\alpha) = 1$. If $P(\alpha) = 0$, then the machine may never stop. For a string α such that $P(\alpha) = 1$, $NT_M(\alpha)$ and $NS_M(\alpha)$ represent the minimum number of steps and the minimum length of tape used in computing $P(\alpha)$. Nondeterministic complexities $NT_M(n)$ and $NS_M(n)$ for the size n are defined in a similar way as before. A problem is said to be solvable in **NP-time** or **NP-space** if it can be solved by a nondeterministic Turing machine M whose time complexity $NT_M(n)$ or space complexity $NS_M(n)$ is bounded by a polynomial function of size n . The class of all problems solvable in NP-time is denoted by **NP**, while that of the problems solvable in polynomial time by ordinary (deterministic) Turing machines is denoted by **P**. Obviously, the class **NP** contains the class **P**. Nevertheless, whether **NP** = **P** or not remains one of the biggest unsolved problems.

A decision problem (Σ, L, P) is **polynomially transformable** to a problem (Σ', L', P') if there exists a mapping h from Σ^+ to Σ'^+ satisfying the following conditions: (1) $h(L) \subseteq L'$; (2) for every string α in L , $P(\alpha) = P'(h(\alpha))$; (3) the mapping h is computed by a Turing machine in polynomial time. If a problem C is polynomially transformable to another problem C' in **P** or **NP**, then the problem C is also in **P** or **NP**. A problem C is said to be **NP-hard** if every problem in **NP** is polynomially transformable to C . It is called **NP-complete** if it is in **NP** and NP-hard. Many problems that have been known empirically to be very hard have recently been shown to be NP-complete [8]. For instance, integer linear programming is

NP-complete. Linear programming has recently been shown to be in **P**. The satisfiability problem of a conjunctive normal form was proven to be NP-complete in 1971 [9].

The **complement** of a decision problem (Σ, L, P) is the problem (Σ, L, P') , where $P'(x) = 1 - P(x)$. If a problem C is in **P**, then its complement C' is also in **P**. When C is in **NP**, it is not guaranteed that C' is in **NP**. The class of all problems whose complements are in **NP** is denoted by **co-NP**. The intersection of **NP** and **co-NP** contains the class **P**, but it is an open problem whether this containment is proper or not. The problem **PRIME**, which asks whether a number N is prime or not, the size n being the number of digits for representing the number N , belongs to both **NP** and **co-NP** [10]. However, it is not known whether it is in **P**.

F. Other Topics

The notion of completeness is introduced in many other classes of problems solvable in linear space, polynomial space, exponential time, etc. Comparisons among these classes have been done, although many problems still remain open. The fundamental conjecture in this field is the nonequality $\mathbf{NP} \neq \mathbf{P}$, the proving of which seems to be extremely difficult. This conjecture can be relativized in ways such that it can be either true or false, so that it is not provable in some formal system.

Some tricks may be used for hard problems. For decision problems, some algorithms have been proposed for guessing the correct answer with high probability [11]. Some algorithms give nearly optimal solutions efficiently, instead of constructing optimal ones over too long a time [7].

A general theory of computational complexity would include investigations on relations among various complexity measures, the complexity hierarchy, and an axiomatic approach independent of any machine models [12].

References

- [1] J. E. Savage, The complexity of computing, Wiley, 1971.
- [2] A. N. Kolmogorov, Three approaches to the quantitative definition of information, Problems of Information Transmission, 1 (1965), 1–7.
- [3] P. Martin-Löf, The definition of random sequences, Information and Control, 9 (1966), 602–619.
- [4] M. J. Fischer and M. O. Rabin, Super-exponential complexity of Presburger arith-

metic, Complexity of Computation, R. M. Karp (ed.), Proc. SIAM-AMS Symposium in Appl. Math., 1974.

[5] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, The design and analysis of computer algorithms, Addison-Wesley, 1974.

[6] D. E. Knuth, The art of computer programming, Addison-Wesley, 1, 1968; 2, 1969; 3, 1973.

[7] E. Horowitz and S. Sahni, Fundamentals of computer algorithms, Computer Science Press, 1978.

[8] M. R. Garey and D. S. Johnson, Computers and intractability, a guide to the theory of NP-completeness, Freeman, 1979.

[9] S. A. Cook, The complexity of theorem-proving procedures, Proc. 3rd Ann. ACM Symposium on Theory of Computing, 1971.

[10] G. L. Miller, Riemann's hypothesis and test for primality, Proc. 7th Ann. ACM Symposium on Theory of Computing, 1975.

[11] J. F. Traub (ed.), Algorithms and complexity: New directions and recent results, Academic Press, 1976.

[12] M. Blum, A machine-independent theory of the complexity of recursive functions, J. Assoc. Comput. Mach., 14 (1967), 322–336.

72 (VII.9) Complex Manifolds

A. Definitions

A Hausdorff topological space X is called a **complex manifold** (or **complex analytic manifold**) of **complex dimension** n if there are given an open covering $\{U_i\}_{i \in I}$ and a family $\{\varphi_i\}_{i \in I}$ of homeomorphisms of U_i onto open sets in the n -dimensional complex affine space \mathbb{C}^n such that in case $U_i \cap U_j \neq \emptyset$, the mapping $\varphi_i \circ \varphi_j^{-1}: \varphi_j(U_i \cap U_j) \rightarrow \varphi_i(U_i \cap U_j)$ is biholomorphic (i.e., $\varphi_i \circ \varphi_j^{-1}$ and its inverse are both holomorphic functions when expressed in terms of coordinate functions in \mathbb{C}^n). We call X the **underlying topological space** of this complex manifold, and we say that an open covering $\{U_i\}_{i \in I}$ and a family $\{\varphi_i\}_{i \in I}$ define a **complex analytic structure** (or simply **complex structure**) on X .

A complex-valued function f defined on an open set U in X is called a **holomorphic function** on U if for any i the function $f \circ \varphi_i^{-1}$ on $\varphi_i(U \cap U_i)$ is holomorphic. When we express the mapping φ_i as $\varphi_i(p) = (z^1(p), \dots, z^n(p))$ on U_i in terms of the coordinates in \mathbb{C}^n , each z^a is a holomorphic function on U_i . We call (z^1, \dots, z^n) a **holomorphic local coordinate system** on U_i . Given two complex manifolds Y, X , a mapping $\varphi: Y \rightarrow X$ is said to be **holo-**

morphic if for any open set U in X and any holomorphic function f on U , $f \circ \varphi$ is holomorphic on $\varphi^{-1}(U) \subset Y$. When a mapping $\varphi: Y \rightarrow X$ is bijective and both φ and φ^{-1} are holomorphic, we say Y and X are **isomorphic** by φ as complex manifolds.

As in the case of \dagger differentiable manifolds of class C^∞ , we can define concepts such as **complex analytic submanifolds**, **holomorphic tangent vectors**, **holomorphic vector fields**, and **holomorphic differential forms of degree k** (or simply **holomorphic k -forms**). **Meromorphic functions** on complex manifolds can also be defined as in the theory of analytic functions of several complex variables (\rightarrow 23 Analytic Spaces D).

Let X be a complex manifold and p a point of X . Take a holomorphic local coordinate system (z^1, \dots, z^n) with center p (i.e., $z^\alpha(p) = 0$ for all α). A holomorphic function defined on a neighborhood of p can be expressed as a holomorphic function in (z^1, \dots, z^n) , hence as a power series in (z^1, \dots, z^n) absolutely convergent in a neighborhood of p . If we denote by $\mathcal{O} = \mathcal{O}_X$ the \dagger sheaf of germs of holomorphic functions on X , the \dagger stalk \mathcal{O}_p of \mathcal{O} at p is isomorphic to the \dagger local ring of convergent power series in n variables z^1, \dots, z^n . At a point p , $(\partial/\partial z^1)_p, \dots, (\partial/\partial z^n)_p$ form a basis of the holomorphic tangent vector space at p . A holomorphic k -form ω defined on a neighborhood of p can be expressed as $\omega = \sum_{i_1 < \dots < i_k} f_{i_1, \dots, i_k} dz^{i_1} \wedge \dots \wedge dz^{i_k}$, where f_{i_1, \dots, i_k} is a holomorphic function for each (i_1, \dots, i_k) .

B. Almost Complex Structures

Let X be a complex manifold, and let $\{U_i, \varphi_i\}_{i \in I}$ be its complex analytic structure, i.e., a covering of X by holomorphic local coordinate systems with $\varphi_i = (z_i^1, \dots, z_i^n)$. Express z_i^α in the form $z_i^\alpha = x_i^\alpha + \sqrt{-1} y_i^\alpha$, where x_i^α and y_i^α are the real and imaginary parts of z_i^α , respectively. Then x_i^α and y_i^α real-valued functions on the open set U_i of X , and the mapping $\psi_i: U \rightarrow \mathbf{R}^{2n}$ defined by $\psi_i(p) = (x_i^1(p), y_i^1(p), \dots, x_i^n(p), y_i^n(p))$ is a homeomorphism of U_i onto an open set of \mathbf{R}^{2n} . This $\{U_i, \psi_i\}_{i \in I}$ defines on X a \dagger differentiable structure of class C^∞ (in fact, a \dagger real analytic structure). Thus a complex manifold of complex dimension n admits canonically a C^∞ -structure of real dimension $2n$. For every point p of X there is a real coordinate system on a neighborhood of p , such as $(x^1, y^1, \dots, x^n, y^n)$, where (z^1, \dots, z^n) ($z^\alpha = x^\alpha + \sqrt{-1} y^\alpha$, $\alpha = 1, \dots, n$) forms a holomorphic coordinate system in X . The real tangent vector space at a point p of X has $\{(\partial/\partial x^1)_p, (\partial/\partial y^1)_p, \dots, (\partial/\partial x^n)_p, (\partial/\partial y^n)_p\}$ as its basis. Define a linear endomorphism J_p by $(\partial/\partial x^\alpha)_p$

$\rightarrow (\partial/\partial y^\alpha)_p$, $(\partial/\partial y^\alpha)_p \rightarrow -(\partial/\partial x^\alpha)_p$ ($\alpha = 1, \dots, n$); then $J_p^2 = -1$, and J_p does not depend on the choice of holomorphic coordinate system at p . Considering J_p as a tensor of type $(1, 1)$, we thus obtain a tensor field J of type $(1, 1)$ of class C^∞ on X , which is called the **tensor field of almost complex structure** induced by the complex structure of X .

More generally, when a real differentiable manifold X is provided with a tensor field J of type $(1, 1)$ of class C^∞ such that $J^2 = -1$ (considering J as a linear transformation of vector fields), we say that X admits an **almost complex structure** or that X is an **almost complex manifold**. In this case, for contravariant vector fields x and y on X , we define a tensor field S of type $(1, 2)$ by $S(x, y) = -[x, y] + [J(x), J(y)] - J([J(x), y]) - J([x, J(y)])$. S is called a **Nijenhuis tensor**. An almost complex structure J is induced by a complex analytic structure if and only if its Nijenhuis tensor S vanishes identically [49]. A differentiable manifold X of dimension $2n$ admits an almost complex structure if and only if the structure group $GL(2n, \mathbf{R})$ of the bundle of \dagger tangent $2n$ -frames of X can be \dagger reduced to $GL(n, \mathbf{C})$. Almost complex manifolds are \dagger orientable.

C. Types of Differential Forms

Let X be a complex manifold, and let (z^1, \dots, z^n) be a holomorphic local coordinate system in a neighborhood of a point p with $z^\alpha = x^\alpha + \sqrt{-1} y^\alpha$ ($\alpha = 1, \dots, n$). On the complexified real tangent vector space $T_p(X) \otimes \mathbf{C}$ at p , we define $\partial/\partial z^\alpha, \partial/\partial \bar{z}^\alpha$ by

$$(\partial/\partial z^\alpha)_p = (1/2) \{ (\partial/\partial x^\alpha)_p - \sqrt{-1} (\partial/\partial y^\alpha)_p \},$$

$$(\partial/\partial \bar{z}^\alpha)_p = (1/2) \{ (\partial/\partial x^\alpha)_p + \sqrt{-1} (\partial/\partial y^\alpha)_p \}.$$

It is easy to see that the operation $\partial/\partial z^\alpha$ on holomorphic functions coincides with that of the holomorphic tangent vector $\partial/\partial z^\alpha$ defined in Section A. A function f is holomorphic at p if and only if $(\partial/\partial \bar{z}^\alpha)f = 0$ ($\alpha = 1, 2, \dots, n$). $T_p(X) \otimes \mathbf{C}$ is the direct sum of the subspace spanned by $\{\partial/\partial z^1, \dots, \partial/\partial z^n\}$ and the subspace spanned by $\{\partial/\partial \bar{z}^1, \dots, \partial/\partial \bar{z}^n\}$. Moreover, this decomposition is independent of the choice of holomorphic coordinate system. Elements of the two subspaces are respectively called **tangent vectors of type $(1, 0)$** and **tangent vectors of type $(0, 1)$** . Similarly, the complexified space of real differentiable 1-forms can be decomposed into the direct sum of two subspaces spanned by $\{dz^1, \dots, dz^n\}$ and $\{d\bar{z}^1, \dots, d\bar{z}^n\}$, where $dz^\alpha = dx^\alpha + \sqrt{-1} dy^\alpha$ and $d\bar{z}^\alpha = dx^\alpha - \sqrt{-1} dy^\alpha$. We say that the elements of the former subspace are of type $(1, 0)$ and those of the latter are of type $(0, 1)$. Thus the space of differential forms of arbitrary degree can be written as

the direct sum of subspaces of type (r, s) . Here the subspace of **differential forms of type (r, s)** has as a basis $\{dz^{\alpha_1} \wedge \dots \wedge dz^{\alpha_r} \wedge d\bar{z}^{\beta_1} \wedge \dots \wedge d\bar{z}^{\beta_s} \mid (1 \leq \alpha_1 < \dots < \alpha_r \leq n, 1 \leq \beta_1 < \dots < \beta_s \leq n)\}$. This decomposition is independent of the choice of a local coordinate system, hence the concept of type can be defined globally on X .

D. $\bar{\partial}$ -Cohomology

In the rest of this article we consider only complex differential forms. For every differential form ω of type (r, s) on X , its \dagger exterior derivative $d\omega$ decomposes into a sum of differential forms of types $(r+1, s)$ and $(r, s+1)$, which we denote by $\partial\omega$ (or $d'\omega$) and $\bar{\partial}\omega$ (or $d''\omega$), respectively. We have $d = \partial + \bar{\partial}$, $(d')^2 = 0$, $(\bar{\partial})^2 = 0$, and $\partial\bar{\partial} + \bar{\partial}\partial = 0$. In terms of a local coordinate system, we write

$$\begin{aligned}\partial\omega &= \sum_{\gamma} \frac{\partial f}{\partial z^{\gamma}} dz^{\gamma} \wedge dz^{\alpha_1} \wedge \dots \\ &\quad \wedge dz^{\alpha_r} \wedge d\bar{z}^{\beta_1} \wedge \dots \wedge d\bar{z}^{\beta_s}, \\ \bar{\partial}\omega &= (-1)^r \sum_{\gamma} \frac{\partial f}{\partial \bar{z}^{\gamma}} dz^{\alpha_1} \wedge \dots \\ &\quad \wedge dz^{\alpha_r} \wedge d\bar{z}^{\gamma} \wedge d\bar{z}^{\beta_1} \wedge \dots \wedge d\bar{z}^{\beta_s},\end{aligned}$$

for $\omega = f dz^{\alpha_1} \wedge \dots \wedge dz^{\alpha_r} \wedge d\bar{z}^{\beta_1} \wedge \dots \wedge d\bar{z}^{\beta_s}$. A differential k -form ω is holomorphic if and only if ω is of type $(k, 0)$ and $\bar{\partial}\omega = 0$.

For the operator $\bar{\partial}$ **Dolbeault's lemma** holds: Let ω be a differential form on a neighborhood U of a point p . If $\bar{\partial}\omega = 0$, there is a neighborhood V of p contained in U and a differential form θ on V such that $\omega = \bar{\partial}\theta$ on V .

Let $A^{(r,s)}$ and Ω^p be the \dagger sheaf of germs of differential forms of type (r, s) and the sheaf of germs of holomorphic p -forms on X , respectively, and let $\Gamma(X, A^{(r,s)})$ be the set of \dagger sections of $A^{(r,s)}$ on X . $\Gamma(X, A^{(r,s)})$ is the set of differential forms of type (r, s) on X , and $\sum_i \Gamma(X, A^{(p,i)})$ forms a \dagger cochain complex with respect to $\bar{\partial}$. This complex is called the **$\bar{\partial}$ -complex** or the **Dolbeault complex**, and its \dagger cohomology groups are called the **$\bar{\partial}$ -cohomology groups** or the **Dolbeault cohomology groups**. The q th cohomology group is denoted by $H^{p,q}(X, \bar{\partial})$. It follows easily from Dolbeault's lemma that $0 \rightarrow \Omega^p \rightarrow A^{(p,0)} \xrightarrow{\bar{\partial}} A^{(p,1)} \rightarrow \dots$ is an \dagger exact sequence of sheaves. From this we get **Dolbeault's theorem**: $H^q(X, \Omega^p) \cong H^{p,q}(X, \bar{\partial})$,

where the left-hand side is the cohomology group with coefficient sheaf Ω^p .

More generally, for any \dagger complex analytic (holomorphic) vector bundle E on X , we can define the $\bar{\partial}$ cohomology groups of the differential forms on X with values in E , and they can be shown to be isomorphic to the coho-

mology groups with coefficient sheaves of germs of holomorphic forms with values in E (\rightarrow 194 Harmonic Integrals E).

E. Analytic Coherent Sheaves

The structure of a complex manifold X is determined by the \dagger sheaf \mathcal{O} of germs of holomorphic functions on X , and \mathcal{O} is a \dagger coherent sheaf (of rings) (**Oka's theorem**). Sheaves of \mathcal{O} -modules are called **analytic sheaves**, and coherent sheaves of \mathcal{O} -modules are called **coherent analytic sheaves**. Many properties of X can be expressed in terms of coherent analytic sheaves and their cohomology groups; some examples appear later in this article (also \rightarrow 366 Riemann-Roch Theorems B).

It is important to know whether an analytic sheaf on a complex manifold is coherent. To this question, not only does Oka's theorem apply but so does **Cartan's theorem**: The \dagger sheaf of ideals defined by an analytic subset of a complex manifold is coherent. (We say that a subset Y of X is an **analytic subset** if it is a closed subset and each point of Y has a neighborhood U such that $U \cap Y$ is the set of common zeros of a finite number of holomorphic functions on U .) Also relevant is **Grauert's theorem**: If $\pi: X \rightarrow Y$ is a \dagger proper holomorphic mapping of complex manifolds (i.e., the inverse image of any compact subset of Y for a holomorphic mapping π is also compact), then for any coherent analytic sheaf F on X its \dagger direct images $R^q \pi_*(F)$ ($q=0, 1, 2, \dots$) are also coherent [16]. (In fact, this theorem holds for analytic spaces; \rightarrow 23 Analytic Spaces.)

For an analytic coherent sheaf F on a \dagger Stein manifold X , we have the following **fundamental theorems of the Stein manifold**. **Theorem A**: $H^0(X, F)$ generates the stalk F_x (as an \mathcal{O}_x -module) at every point x of X . **Theorem B**: $H^q(X, F) = 0$ for all $q > 0$. Conversely, a Stein manifold X is characterized by the following property: For any coherent analytic sheaf F of ideals of \mathcal{O} , $H^1(X, F) = 0$. If a complex manifold X is compact and F is a coherent analytic sheaf on X , then $H^q(X, F)$ is a complex vector space of finite dimension. If X is an open submanifold of another complex manifold and its closure is compact, then $H^q(X, F)$ is finite-dimensional for some q that depends on various properties (convexity or concavity) of the boundary of X [3].

Let E be a \dagger complex analytic (holomorphic) vector bundle on a complex manifold X of dimension n , and let E^* be the \dagger dual vector bundle of E . Then $H^q(X, \mathcal{O}^p(E))$ and $H^{n-q}_*(X, \mathcal{O}^{n-p}(E^*))$ (where H_* denotes the cohomology group with compact support) are dual as topological vector spaces under suit-

able conditions. The duality is given by the integration on X of the exterior products of the differential forms representing the respective elements of the cohomology groups. The duality holds, for example, when $\dim H^q(X, \mathcal{O}^p(E)) < \infty$. If X is compact, we need not distinguish H_* from H (**Serre's duality theorem**) [50].

F. Compact Complex Manifolds

On a compact and connected complex manifold X , there are no holomorphic functions except constants (by the †maximum principle of holomorphic functions). The field $K(X)$ of meromorphic functions on X is finitely generated over the complex number field, and its †transcendence degree d does not exceed the complex dimension n of X . d is said to be the **algebraic dimension** of X and is denoted by $a(X)$. For elements of $K(X)$, functional independence and algebraic independence are equivalent [52]. When $n=1$, X is a compact †Riemann surface and the classical theory of algebraic functions shows that $K(X)$ is an †algebraic function field of one variable and X is a †projective algebraic variety. When $n=2$, $a(X)=2, 1$, and 0 can all occur. If $a(X)=2$, X is a projective †algebraic surface (**Chow-Kodaira theorem**). If $a(X)=1$, there exist an †algebraic curve Δ and a surjective holomorphic mapping $\varphi: X \rightarrow \Delta$ such that $K(\Delta)$ is isomorphic to $K(X)$ under φ^* , and $\varphi^{-1}(x)$ is an †elliptic curve for all but a finite number of $x \in \Delta$. K. Kodaira investigated the structure of compact complex surfaces in detail [34, III].

On a compact complex manifold X , the free Abelian group generated by the set of irreducible analytic subsets of codimension 1 is called the **divisor group** of X , and an element of it is called a **divisor** of X . For an analytic subset Y of codimension 1, the sheaf of ideals $\mathfrak{I}(Y)$ defined by Y is a sheaf of locally principal ideals of \mathcal{O} . For a divisor $D = \sum a_\alpha Y_\alpha$, the sheaf of locally principal fractional ideals $\mathfrak{I}(D) = \prod_\alpha \mathfrak{I}(Y_\alpha)^{a_\alpha}$ is called the **sheaf of ideals of D** . The set of nonzero coherent sheaves of locally principal fractional ideals corresponds bijectively to the set of divisors. An element $f \neq 0$ of $K(X)$ generates a sheaf of principal fractional ideals and therefore defines a divisor, which is denoted by (f) . The divisor group has an ordering defined by $D = \sum a_\alpha Y_\alpha > 0$ if and only if all $a_\alpha \geq 0$, under which it becomes an †ordered group. For a divisor D , let

$$L(D) = \{f \in K(X) | f \neq 0, (f) + D \geq 0\} \cup \{0\}.$$

Then $L(D)$ is a \mathbb{C} -module of finite dimension. This submodule of $K(X)$ is easy to handle and exhibits various analytic properties of D . The

†Riemann-Roch theorem is used to calculate the dimension of $L(D)$ in terms of other factors (\rightarrow 366 Riemann-Roch Theorems C). Here is an example of how $L(D)$ exhibits a property of D : We call two divisors D and D' **linearly equivalent** if there is a $0 \neq F \in K(X)$ such that $(F) = D - D'$. This is an equivalence relation finer than †homological equivalence. If D and D' are linearly equivalent, then $L(D)$ and $L(D')$ are isomorphic by the mapping $L(D) \ni f \rightarrow fF \in L(D')$; therefore $\dim L(D) = \dim L(D')$. (The latter equation does not follow from the homological equivalence.) A holomorphic vector bundle with fiber \mathbb{C} and structure group \mathbb{C}^* is called a **complex line bundle**. For the sheaf of ideals $\mathfrak{I}(D)$ of a divisor D , we can take a suitable open covering $\{U_j\}$ of X such that for each U_j there is an $R_j \in \Gamma(U_j, \mathfrak{I}(D))$ which generates $\mathfrak{I}(D)_x$ for any $x \in U_j$. Also $g_{jk} = R_j/R_k$ is a holomorphic function nowhere vanishing on $U_j \cap U_k$. With $\{g_{jk}\}$ as the system of †coordinate transformations, we define the **complex line bundle determined by D** and denote it by $[D]$. It is easy to see that $[D]$ is independent of the choice of $\{U_j\}$ or $\{R_j\}$. Moreover, $[D]$ is determined only by the linear equivalence class of D . If we denote by $\mathcal{O}([D])$ the sheaf of germs of holomorphic sections of $[D]$, the mapping $H^0(X, \mathcal{O}([D])) \ni \varphi = \{\varphi_j\} \rightarrow f = \varphi_j/R_j = \varphi_k/R_k \in L(D)$ is an isomorphism of these modules. On an algebraic variety in a projective space, any complex line bundle comes from a divisor (i.e., it can be expressed in the form $[D]$ for some divisor D) [35, 51], but this is not necessarily true on general compact complex manifolds. However, the importance of complex line bundles in the theory of complex manifolds lies in the relation $L(D) \cong H^0(X, \mathcal{O}([D]))$, which replaces “things with poles” with “things holomorphic.”

Any analytic subvariety X of the projective space \mathbb{P}^N is an algebraic variety (**Chow's theorem**) [51]. Suppose that X is an analytic submanifold of \mathbb{P}^N , and let Y be a general †hyperplane section of X . Then Y is a divisor on X , and the †Chern class of the complex line bundle $[Y]$ corresponds to the canonical Hodge metric on X (\rightarrow 232 Kähler Manifolds D). When $[Y]$ is represented by the system of coordinate transformations $\{g_{jk}\}$ with respect to an open covering $\{U_i\}$, we can associate with a coherent analytic sheaf F sheaves $F(n)$ ($n=0, \pm 1, \pm 2, \dots$) as follows. Denoting by F_j the restriction of F to U_j , we glue F_j and F_k together on $U_j \cap U_k$ with the relation $f_j \sim f_k \Leftrightarrow f_j = g_{kj}^{-n} f_k$ (where $f_j \in F_j \subset F$, $f_k \in F_k \subset F$) and obtain a sheaf (denoted by $F(n)$) that is locally isomorphic to F . The following theorems for $F(n)$ hold. For each coherent analytic sheaf F there exists an integer n_0 such that for any $n \geq n_0$ the following **fundamental theorems A, B of projec-**

tive algebraic varieties hold [17]. **Theorem A:** $\Gamma(X, F(n))$ generates F_x (as an \mathcal{O}_x -module) for every $x \in X$. **Theorem B:** $H^q(X, F(n)) = 0$ for all $q > 0$. This means that if we permit "poles" on Y of sufficiently high order, then F has sufficiently many sections and the higher cohomology groups vanish.

On a (nonsingular) algebraic variety X in \mathbf{P}^N we have the sheaf \mathcal{O} of germs of holomorphic functions (the \dagger structure sheaf as a complex manifold) and the sheaf \mathcal{O}_{alg} of germs of holomorphic rational functions (the structure sheaf as an algebraic variety). Therefore we have two kinds of coherent sheaves, coherent analytic sheaves and **coherent algebraic sheaves**. In fact, the cohomology theories derived from them are isomorphic. More precisely, for any coherent algebraic sheaf F , $\tilde{F} = F \otimes_{\mathcal{O}_{\text{alg}}} \mathcal{O}$ is a coherent analytic sheaf. The correspondence $F \rightarrow \tilde{F}$ gives an equivalence between the \dagger category of coherent algebraic sheaves and that of coherent analytic sheaves, thus giving an isomorphism of their cohomology groups. In other words, as far as the properties that can be expressed by cohomology-theoretic terms of coherent sheaves are concerned, there is no difference between the analytic and algebraic theories of projective algebraic varieties [51].

G. Deformations of Complex Structures

The deformation theory of complex structures was initiated by Kodaira and Spencer [36] in order to explain various phenomena in the theory of (compact) complex manifolds. A triple (X, π, S) is called a **family of compact complex manifolds** if X and S are connected \dagger analytic spaces and π is a \dagger proper holomorphic mapping of X onto S such that (i) it is smooth, i.e., is locally identified with the projection $S' \times U \rightarrow S'$, where S' and U are open in S and \mathbf{C}^n , respectively, and (ii) every fiber $V_s = \pi^{-1}(s)$ of π is connected. V_s is then a compact complex manifold. We sometimes write $\{V_s\}_{s \in S}$ instead of (X, π, S) . S is called the **parameter space of the family**. Take a point $o \in S$ and fix it. We say that V_s is a **deformation** of V_o for any $s \in S$. There are a neighborhood S' of o in S and a diffeomorphism $\pi^{-1}(S') \rightarrow S' \times V_o$ (Kuranishi [39]). Thus every $V_s, s \in S'$, determines a \dagger complex structure on a fixed differentiable manifold. We say that the complex structure of $V_s, s \in S'$, is a **deformation** of that of V_o .

Now $(X, \pi, S) = \{V_s\}_{s \in S}$ is said to be **complete at o** if for any family $(Y, \mu, T) = \{W_t\}_{t \in T}$ with a point $o' \in T$ and a holomorphic isomorphism $\iota: W_{o'} \rightarrow V_o$, there is a neighborhood T' of o' in T and holomorphic mappings $f: T' \rightarrow S$ and $h: \mu^{-1}(T') \rightarrow X$ such that (i) $f\mu = \pi h$, (ii) $f(o') = o$,

and (iii) $h = \iota$ on $W_{o'}$. In this case, if T' is sufficiently small, then h induces a holomorphic isomorphism $h_t: W_t \rightarrow V_{f(t)}$, for any $t \in T'$. Hence $\{V_s\}_{s \in S}$ contains all small deformations of V_o . On the other hand, $\{V_s\}_{s \in S}$ is said to be **effectively parametrized at o** if the **Kodaira-Spencer mapping (Kodaira-Spencer map)** $\rho_o: T_o S \rightarrow H^1(V_o, \Theta)$ is injective, where $T_o S$ is the Zariski tangent space to S at o and Θ is the sheaf of germs of holomorphic vector fields on V_o . Here, the linear mapping ρ_o is defined by $\rho_o(\partial/\partial s) = \{(\partial g_{jk}/\partial s)_{(z_k, o)}\}$, where $\partial/\partial s \in T_o S$ and g_{jk} are the coordinate transformations $(z_j, s) = (g_{jk}(z_k, s), s)$, using the smoothness of π . $\rho_o(\partial/\partial s)$ is called the **infinitesimal deformation to the direction $\partial/\partial s$** .

Kuranishi's fundamental theorem [39, 40] states: For any compact complex manifold V , there exists a family $\{V_s\}_{s \in S}$ with a point $o \in S$ such that (i) it is complete at every point of S , (ii) it is effectively parametrized at o , and (iii) $V_o = V$.

The parameter space S in the theorem is called the **Kuranishi space** or the **local moduli space of V** . It is given as the zeros of a holomorphic mapping $f: U \rightarrow H^2(V, \Theta)$ with $f(0) = 0$, where U is a neighborhood of 0 in $H^1(V, \Theta)$. Hence (1) if $H^2(V, \Theta) = 0$, then $S = U$ is nonsingular (Kodaira, Nirenberg, and Spencer [37]; (2) if $H^1(V, \Theta) = 0$, then $S = \{0\}$, i.e., one point (e.g., $V = \mathbf{P}^n(\mathbf{C})$, the complex projective space).

Kuranishi's theorem was generalized to a compact analytic space V by Grauert [18] and A. Douady [8].

There are Kuranishi-type theorems in the deformation theory of other objects: (1) compact analytic subvarieties of an analytic space (Douady [9]; → 23 Analytic Spaces G), (2) holomorphic mappings (Y. Miyajima [41]), (3) germs of analytic spaces with isolated singularities [7], etc.

So far, only the local theory of deformations has been developed. The global theory is not yet in a satisfactory state. Its final purpose is to construct moduli spaces and to understand them. For a compact differentiable manifold V , we denote by $M(V)$ the set of all isomorphism classes of complex structures on V . It is difficult in general to determine the set $M(V)$. As yet unsolved problems are: (1) Is $M(S^6)$ nonempty? (2) Does $M(\mathbf{P}^n(\mathbf{C}))$ ($n \geq 3$) consist of one point? ($M(\mathbf{P}^2(\mathbf{C}))$ is known to consist of one point (S. T. Yau; → 232 Kahler Manifolds C).

If $M(V)$ has a reasonable structure (e.g., an analytic space structure) and a universal property, then we call it the **moduli space** (→ 11 Algebraic Functions F, 16 Algebraic Varieties W). Kodaira [34, III] constructed the moduli space of Hopf surfaces. Only a few examples of moduli spaces are known. It is to be noted that moduli spaces cannot in general exist.

One of the reasons for this is that there may exist **jumpings of structures** (Kodaira and Spencer [36]).

As for algebraic manifolds (varieties), there are two known methods for the moduli problem: (1) Griffiths' period mapping (\rightarrow 16 Algebraic Varieties V) and (2) Mumford's geometric invariant theory (\rightarrow 16 Algebraic Varieties V, W).

H. Monoidal Transformations

Let Y be an analytic subspace of a complex manifold X , which is defined by a nonzero coherent sheaf of ideals \mathfrak{I} . Then X has an open covering such that for each member U of the covering, there are elements $\varphi_1, \dots, \varphi_m \in \Gamma(U, \mathfrak{I})$ that generate the stalks \mathfrak{I}_x at all $x \in U$. Let W' be the graph of the holomorphic mapping $x \mapsto (\varphi_1(x) : \varphi_2(x) : \dots : \varphi_m(x))$ from $U - U \cap Y$ to \mathbf{P}^{m-1} , and denote by W the closure of W' in $U \times \mathbf{P}^{m-1}$. Then W is an analytic space, possibly with singularities, which is independent of the choice of the generators $\{\varphi_i\}$ and is determined uniquely by U and \mathfrak{I} . Therefore all the W 's can be glued together to form an analytic space \tilde{X} and to determine a holomorphic mapping $p: \tilde{X} \rightarrow X$. We call p the **monoidal transformation** (or **blowing-up**) of X with center \mathfrak{I} or with center Y . When Y is a point, p is also called a **locally quadratic transformation** or a **σ -process**. When Y is an analytic submanifold, \tilde{X} also is a manifold, $\tilde{Y} = p^{-1}(Y)$ is a nonsingular divisor on \tilde{X} , and $p^{-1}(y) \cong \mathbf{P}^{k-1}$ for all $y \in Y$, where k is the codimension of Y . Moreover, the line bundle $[\tilde{Y}]$ restricted to each fiber $p^{-1}(y)$ is isomorphic to the Hopf bundle L on \mathbf{P}^{k-1} , i.e., the line bundle associated to the natural \mathbf{C}^* bundle $\mathbf{C}^k \rightarrow \mathbf{P}^{k-1}$. Conversely, suppose that we are given $\tilde{Y} \subseteq \tilde{X}$ and a holomorphic mapping $p_0: \tilde{Y} \rightarrow Y$ with the property as above so that, in particular, $p_0^{-1}(y) \cong \mathbf{P}^{k-1}$ and $[\tilde{Y}] \cong L$ on $p_0^{-1}(y)$ for all $y \in Y$. Then there exists a complex manifold containing Y as a submanifold such that \tilde{X} is obtained by the monoidal transformation of X with center Y as above (S. Nakano [47]). Let $f: X \rightarrow Y$ be a proper modification of complex manifolds (or analytic spaces). Then there exists a proper holomorphic mapping $h: Y' \rightarrow Y$ such that $f^{-1} \circ h: Y' \rightarrow X$ is holomorphic, where over any relatively compact subdomain of Y , h is obtained by a finite succession of monoidal transformations with nonsingular centers (Hironaka). The result is called **Chow's lemma**.

I. Fiber Spaces

A triple of compact complex manifolds V , W , and a surjective holomorphic mapping $f:$

$V \rightarrow W$ is called a **fiber space** if general fibers $f^{-1}(w)$ are irreducible. In addition, if both V and W are algebraic, it is called an **algebraic fiber space**.

For a compact complex manifold V , letting K_V denote the canonical line bundle of V , define a subfield $\mathfrak{K}(V)$ to be $\{\omega_1/\omega_2, \text{ where } \omega_1, \omega_2 \in H^0(V, \Omega) \text{ for some } m > 0\}$. The transcendental degree of \mathfrak{K} over \mathbf{C} is denoted by $\kappa(V)$, which we call the **Kodaira dimension** of V .

If $\mathfrak{K} = 0$, then define $\kappa(V) = -\infty$. If $\kappa(V) = \dim V$, then V is said to be of **general type**, and there exist a projective manifold X and a bimeromorphic holomorphic mapping $\mu: X \rightarrow V$. If $\kappa(V) \geq 0$, there exists a fiber space $f: V^* \rightarrow W$ such that (1) V^* is bimeromorphically equivalent to V , (2) $\dim W = \kappa(V)$, and (3) some general fibers $f^{-1}(w)$ satisfy $\kappa(f^{-1}(w)) = 0$. In general, for a fiber space $f: V \rightarrow W$, we have $\kappa(V) \leq \kappa(f^{-1}(w)) + \dim W$, where $f^{-1}(w)$ is a general fiber. Moreover, if it is algebraic, an inequality of the form $\kappa(V) \geq \kappa(f^{-1}(w)) + \kappa(W)$ is called the **conjecture** C_n , n being $\dim V$. C_n has been verified by K. Ueno, E. Viehweg, T. Fujita, and Y. Kawamata [13, 14, 32, 33, 54–57] in the following cases: (1) $n \leq 3$, (2) when general fibers are curves, (3) $\dim w = 1$, $\kappa(w) = 1$, (4) $\kappa(V) \geq 0$, and W is of general types, etc. By using the results of case (4), Kawamata proves that an algebraic compact complex manifold V is birationally equivalent to an Abelian variety if and only if $\kappa(V) = 0$ and the irregularity of V equals the dimension of V [32].

J. Analytic Surfaces

In what follows, an analytic surface means a 2-dimensional compact complex manifold. For an analytic surface S , an exceptional curve on S and a (relatively) minimal model, etc., are defined with respect to bimeromorphic mappings in analogy with the corresponding concepts for an algebraic surface (\rightarrow 15 Algebraic Surfaces). Let $C \subset S$ be an irreducible curve on S . Then there exists a holomorphic mapping φ from S onto another surface S' such that $\varphi(C)$ is a point and such that φ induces the isomorphism $S - C \xrightarrow{\sim} S' - \varphi(C)$ if and only if $C^2 = -1$ and C is a nonsingular rational curve (Grauert [17]). S has a minimal model if and only if S is not a ruled surface (Kodaira [34, III]). The irregularity $q = h^{0,1}$, the geometric genus p_g , i -genus P_i , etc., are also defined in the same way as in the case of algebraic surfaces. Note that, in general, $h^{0,1} \neq h^{1,0}$. The Riemann-Roch theorem and M. Noether's formula are valid also for an analytic surface (Atiyah and Singer; \rightarrow 366 Riemann-Roch Theorems C).

K. Classification of Surfaces

The classification of analytic surfaces by the aid of their numerical invariants was completed by Kodaira and includes as a special case Enriques's classification of algebraic surfaces [34, III]. By an **elliptic surface** we mean a surface from which there exists a surjective holomorphic mapping φ onto an algebraic curve Δ , such that for a general point p on Δ , $\varphi^{-1}(p)$ is an irreducible nonsingular elliptic curve. If $a(S)=1$ or $\kappa(S)=1$, S has a unique structure as an elliptic surface. The image by φ of the points on S at which φ is not of maximal rank is a finite subset $\{a_1, \dots, a_r\}$ of Δ . Let t_i be a local coordinate on Δ around a_i with $t_i(a_i)=0$. We call a singular fiber of φ the divisor on S defined by $\{t_i \circ \varphi = 0\}$. The structure and the construction of singular fibers of elliptic surfaces have been completely determined by Kodaira. By an elliptic surface of general type we mean a surface with Kodaira dimension 1. If $\kappa(S)=2$, S is projective algebraic and is called a **surface of general type**. If $a(S)=0$, then there exists only a finite number of irreducible curves on S . By a **Hopf surface** we mean a surface whose universal covering is $\mathbb{C}^2 - (0,0)$. If an analytic surface S is homeomorphic to $S^1 \times S^3$, S is a Hopf surface. Let $b_v(S)$ be the v th Betti number of S . If $a(S)=0$, $b_1(S)=1$, $b_2(S)=0$, and S contains a curve, then S is a Hopf surface (Kodaira). By a surface of class VII_0 we mean a minimal surface S with $b_1(S)=1$. M. Inoue [26] constructed three families of surfaces of class VII_0 with $b_2=0$ which contain no curves. These are $S_M^{(\pm)}$ ($M \in \text{SL}(3, \mathbb{Z})$) and $S_{N,p,q,r,t}^{(+)}$ and $S_{N,p,q,r,t}^{(-)}$ ($N \in \text{SL}(2, \mathbb{Z})$, $p, q, r \in \mathbb{Z}$, $t \in \mathbb{C}$) which have $H \times \mathbb{C}$ as their universal covering surfaces, H being the complex upper half-plane. These have a line bundle L such that $H^0(\Omega^1 \otimes \mathcal{O}(L))=0$. This property characterizes these Inoue surfaces among VII_0 surfaces with $b_2=0$ which contain no curves. A couple of new surfaces of class VII_0 with $b_2>0$ were constructed and studied by Inoue, Ma, Kato, T. Oda, I. Nakamura, and I. Enoki [10, 27–29, 46]. Some of these have close connections with cusp singularities of Hilbert modular surfaces, torus embeddings, and global spherical shells. Enoki's surface is denoted by $S_{n,\alpha,t}$ ($n>0$, $0<|\alpha|<1$, $t \in \mathbb{C}^n$) and has the following properties: (1) class VII_0 and $b_2=n$, (2) there exists a connected curve D with $D^2=0$, (3) $S_{n,\alpha,t} - D$ is an affine bundle over an elliptic curve. Properties (1) and (2) characterize $S_{n,\alpha,t}$ (Enoki).

Let K_S denote the canonical line bundle of an analytic surface S . An analytic surface S is said to be a **K3 surface**, if K_S is trivial, i.e., if there exists a nonvanishing holomorphic 2-

form ω and if $q=0$. K3 surfaces are simply connected and are deformations of a nonsingular quartic surface in \mathbb{P}^3 (Kodaira). We shall define the period mapping of K3 surfaces. Let L be a free Abelian group of rank 22 with the pairing $\langle \rangle$ which is the direct sum of the two copies of $-E_8$ and three copies of $U = \mathbb{Z}e_1 + \mathbb{Z}e_2$ with $\langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = 0$, $\langle e_1, e_2 \rangle = \langle e_2, e_1 \rangle = 1$, where $-E_8$ denotes the lattice with the pairing corresponding to the Dynkin diagram E_8 with the opposite signs. $H^2(S, \mathbb{Z})$ with the intersection form is such a pair, consisting of S and $\langle \rangle$. A **marked K3 surface** is defined to be a pair (S, ψ) , where S is a K3 surface and $\psi: H^2(S, \mathbb{Z}) \rightarrow L$ is an isomorphism preserving $\langle \rangle$. The class $[\omega]$ of ω is a base of $H^{2,0}$ and satisfies $\langle [\omega], [\omega] \rangle = 0$ and $\langle [\omega], [\bar{\omega}] \rangle > 0$. Let $\mathbf{P}(L_{\mathbb{C}})$ denote the 21-dimensional projective space associated with $L_{\mathbb{C}} = L \otimes_{\mathbb{Z}} \mathbb{C}$. $\mathcal{D} = \{(a) \in \mathbf{P}(L_{\mathbb{C}}) | \langle a, a \rangle = 0, \langle a, \bar{a} \rangle > 0\}$ is an open set of the quadric $\mathcal{Q}_{\mathbb{C}} = \{(a) \in \mathbf{P}(L_{\mathbb{C}}) | \langle a, a \rangle = 0\}$. \mathcal{D} is the set of all Hodge structures on $L_{\mathbb{C}}$. The **period** of a marked K3 surface (S, ψ) is defined to be $(\psi_{\mathbb{C}}[\omega])$. Every point of \mathcal{D} is the period of some marked Kähler K3 surface. If the periods of two marked Kähler K3 surfaces (S, ψ) and (S', ψ') , coincide then S is isomorphic to S' .

An analytic surface with $p_g = q = 0$ and $K_S^{\otimes 2} \cong 0$ is called an **Enriques surface**, which has a K3 surface as its universal covering. An Enriques surface is an algebraic elliptic surface. An analytic surface with $q=1$ and $K_S^{\otimes 12} \cong 0$ is said to be a **hyperelliptic surface**, which has an Abelian surface as an unramified covering. It is an algebraic surface and is an elliptic bundle over an elliptic curve. The classification of minimal surfaces is given in Table 1. The following relations hold among these invariants: Let b^+ (b^-) be the number of positive (negative) eigenvalues (counted with multiplicities) of the intersection matrix on $H^2(S, \mathbb{R})$ and c_i be the i th Chern class of S . Then

(1) $b^+ - b^- = \frac{1}{3}(c_1^2 - 2c_2)$ (**Hirzebruch signature theorem**);

(2) if b_1 is even, $q = h^{1,0} = \frac{1}{2}b_1$ and $b^+ = 2p_g + 1$;

(3) if b_1 is odd, $q = h^{1,0} + 1 = \frac{1}{2}(b_1 + 1)$ and $b^+ = 2p_g$.

Let c_1^2 and c_2 denote the Chern numbers of an analytic surface S , i.e., $c_1^2 = (K_S^2)$ and c_2 is the Euler number of S . Then $3c_2 \geq c_1^2$ [43], and if equality holds, then K_S is ample (Y. Miyaoka) (\rightarrow 232 Kähler Manifolds C). If S is a surface of general type which is minimal, then $c_1^2 > 0$, the bigenus $P_2 \geq 2$ and the m -genus P_m is $m(m-1)c_1^2/2 + 1 - q + p_g$ for $m \geq 2$. For each $m \geq 5$, the m th canonical mapping of S is a birational holomorphic mapping onto its image (Kodaira, E. Bombieri [5]). In general, $c_1^2 \geq 2p_g - 4$, and if the canonical mapping, i.e., the

rational mapping associated with K_S , is birational, then $c_1^2 \geq 3p_g - 7$. For certain kinds of surfaces of general type, E. Horikawa succeeded in determining completely the structure of surfaces obtained as deformations of a given surface. Among others, every minimal surface with $p_g = 4$, $q = 0$, and $c_1^2 = 5$ is a deformation of a nonsingular quintic surface [22, 23].

Table 1. Classification of Minimal Surfaces

κ	p_g	P_{12}	q	b_1	Structure
2		> 0			algebraic surface of general type
1					elliptic surface of general type
0	1	1	2	4	complex torus
	1	1	2	3	elliptic surface with a trivial canonical bundle
	0	1	1	2	hyperelliptic surface
	0	1	1	1	elliptic surface belonging to class VII_0
	1	1	0	0	K3 surface
	0	1	0	0	Enriques surface
$-\infty$	0	0	0	0	rational surface
			≥ 1	$2q$	ruled surface of genus q
			1	1	surface of class VII_0

For $a, b > 0$, let Σ_{ab} be the set of isomorphism classes of all minimal surfaces with $c_1^2 = a$ and $1 - q + p_g = b$. D. Gieseker [15] proved the existence of N and constructed the mapping $h: \Sigma_{ab} \rightarrow \mathbf{P}^N$ so that the following hold: (1) h is injective; (2) if $f: X \rightarrow Y$ is a smooth holomorphic mapping such that all $h^{-1}(y)$ represent some classes $\in \Sigma_{ab}$, then the set-theoretic mapping $f: Y \rightarrow \mathbf{P}^N$ induced from h is a morphism of schemes; (3) $h(\Sigma_{ab})$ is a locally closed subvariety of \mathbf{P}^N . Hence, $h(\Sigma_{ab})$ is a moduli variety of surfaces of general type with $c_1^2 = a$ and $1 - q + p_g = b$.

For each m with $1 \leq m \leq 9$, there exists a minimal surface of general type with $p_g = q = 0$ and $c_1^2 = m$.

If there exist a compact analytic surface S and a curve C on S such that $S - C$ is biholomorphic to a complex manifold M , then we say that S is a **compactification** of M with boundary C . Every compactification of \mathbf{C}^2 is a rational surface (Kodaira, J. Morrow [45]). Every compactification of $\mathbf{C} \times (\mathbf{C} - \{0\})$ is also rational (T. Ueda [53]). However, all compac-

tifications of $(\mathbf{C} - \{0\})^2$ are rational surfaces, certain kinds of Hopf surfaces, or \mathbf{P}^1 -bundles over elliptic curves as constructed by Serre (Ueda).

References

- [1] T. Akahori, Complex analytic construction of the Kuranishi family on a normal strongly pseudo-complex manifold. Publ. Res. Inst. Math. Sci., 14 (1978), 789–847.
- [2] I. Amemiya, K. Masuda, and K. Shiga, Lie algebras of differential operators, Osaka J. Math., 12 (1975), 139–172.
- [3] A. Andreotti and H. Grauert, Théorèmes de finitude pour la cohomologie des espaces complexes, Bull. Soc. Math. France, 90 (1962), 193–259.
- [4] A. Blanchard, Sur les variétés analytiques complexes, Ann. Sci. École Norm. Sup., 73 (1956), 157–202.
- [5] E. Bombieri, Canonical models of surfaces of general type, Publ. Math. Inst. HES, 42 (1973), 171–219.
- [6] M. Commichau, Deformation kompakter komplexer Mannigfaltigkeiten, Math. Ann., 213 (1975), 43–96.
- [7] I. F. Donin, Complete families of deformations of germs of complex spaces (in Russian), Mat. USSR Sb., 18 (1972), 397–406.
- [8] A. Douady, Le problème des modules pour les variétés analytiques complexes, Sémin. Bourbaki, 1964/65, no. 277.
- [9] A. Douady, Le problème des modules locaux pour les espaces \mathbf{C} -analytiques compacts, Ann. Sci. École Norm. Sup., sér. IV, 4 (1974), 569–602.
- [10] I. Enoki, On surfaces of class VII_0 with curves, Proc. Japan Acad., 56 (1980), 275–279.
- [11] O. Forster and K. Knorr, Konstruktion verseller Familien kompakter komplexer Räume, Lecture notes in math. 705, Springer, 1979.
- [12] A. Fujiki, On the blowing down of analytic spaces, Publ. Math. Res. Inst. Math. Sci., 10 (1975), 473–507.
- [13] T. Fujita, On Kähler fiber spaces over curves, J. Math. Soc. Japan, 30 (1978), 779–794.
- [14] T. Fujita, The sheaf of relative canonical forms of a Kähler fiber space over a curve, Proc. Japan Acad., 54 (1978), 183–184.
- [15] D. Gieseker, Global moduli for surfaces of general type, Inventiones Math., 43 (1977), 233–282.
- [16] H. Grauert, Ein Theorem der analytischen Garbentheorie und die Modulräume komplexer Strukturen, Publ. Math. Inst. HES, 5 (1960).
- [17] H. Grauert, Über Modifikationen und

- exzeptionelle analytische Mengen, *Math. Ann.*, 146 (1962), 331–368.
- [18] H. Grauert, Der Satz von Kuranishi für kompakte komplexe Räume, *Inventiones Math.*, 25 (1974), 107–142.
- [19] F. Hirzebruch, Hilbert modular surfaces, *Enseignement Math.*, 19 (1973), 183–281.
- [20] F. Hirzebruch and D. Zagier, Classification of Hilbert modular surfaces, *Complex Analysis and Algebraic Geometry in honor of Kodaira*, Iwanami Shoten (1977), 137–150.
- [21] E. Horikawa, On deformations of holomorphic maps. I, II, *J. Math. Soc. Japan*, 25 (1973), 372–396; 26 (1974), 647–666. III, *Math. Ann.*, 222 (1976), 275–282.
- [22] E. Horikawa, On deformations of quintic surfaces, *Inventiones Math.*, 31 (1975), 43–85.
- [23] E. Horikawa, Algebraic surfaces of general type with small C_2^2 . I, *Ann. Math.*, (2) 104 (1976), 357–387. II, III, IV, *Inventiones Math.*, 37 (1976), 121–155; 47 (1978), 209–248; 50 (1979), 103–128.
- [24] E. Horikawa, Surjectivity of the period map of K3 surfaces of degree 2, *Math. Ann.*, 228 (1977), 113–146.
- [25] S. Iitaka, On D -dimensions of algebraic varieties, *J. Math. Soc. Japan*, 23 (1971), 356–373.
- [26] M. Inoue, On surfaces of class VII_0 , *Inventiones Math.*, 24 (1974), 269–310.
- [27] M. Inoue, New surfaces with no meromorphic functions, *Proc. Int. Congr. Math.*, Vancouver, 1974, vol. 1, 423–426.
- [28] M. Inoue, New surfaces with no meromorphic functions, *Complex Analysis and Algebraic Geometry II*, Iwanami Shoten and Cambridge Univ. Press, 1977, 91–106.
- [29] M. Kato, Complex manifolds containing global spherical shells, *Proc. Int. Symp. Algebraic Geom.* Kyoto, 1977, 45–84.
- [30] S. Kawai, On compact complex analytic manifolds of complex dimension 3. I, II, *J. Math. Soc. Jap.*, 17 (1965), 438–442; 21 (1969), 604–616.
- [31] Y. Kawamata, On deformations of compactifiable complex manifolds, *Math. Ann.*, 235 (1978), 247–265.
- [32] Y. Kawamata, Characterization of abelian varieties, *Compositio Math.*, 43 (1981), 253–276.
- [33] Y. Kawamata, Kodaira dimension of certain algebraic fiber spaces, *J. Fac. Sci. Univ. Tokyo*, sec. IA, 30 (1983), I-24.
- [34] K. Kodaira, Collected works I–III, Iwanami, 1975.
- [35] K. Kodaira and D. C. Spencer, Groups of complex line bundles over compact Kähler varieties, Divisor class groups on algebraic varieties, *Proc. Nat. Acad. Sci. US*, 39 (1953), 868–877.
- [36] K. Kodaira and D. C. Spencer, On de-

- formations of complex analytic structures: I, II, *Ann. Math.*, (2) 67 (1958), 328–466; III, *ibid.*, 71 (1960), 43–76.
- [37] K. Kodaira, L. Nirenberg, and D. C. Spencer, On the existence of deformations of complex analytic structures, *Ann. Math.*, (2) 68 (1958), 450–459.
- [38] R. S. Kulkarni, On the principle of uniformization, *J. Differential Geom.*, 13 (1978), 109–138.
- [39] M. Kuranishi, On the locally complete families of complex analytic structures, *Ann. Math.*, (2) 75 (1962), 536–577.
- [40] M. Kuranishi, New proof for the existence of locally complete families of complex structures, *Proc. Conf. Complex Analysis*, Minneapolis, 1964, Springer, 1965, 142–154.
- [41] K. Miyajima, On the existence of Kuranishi family for deformations of holomorphic maps, *Sci. Rep. Kagoshima Univ.*, 27 (1978), 43–76.
- [42] Y. Miyaoka, Kähler metrics on elliptic surfaces, *Proc. Japan Acad.*, 50 (1974), 533–536.
- [43] Y. Miyaoka, On the Chern numbers of surfaces of general type, *Inventiones Math.*, 42 (1977), 225–237.
- [44] B. G. Moishezon, The algebraic analog of compact complex spaces with a sufficiently large field of meromorphic functions I, II, III, *Math. USSR-Izv.*, 3 (1969), 167–226; 305–343; 447–513.
- [45] J. Morrow, Minimal normal compactification of C^2 , *Rice Univ. Studies*, 59, no. 1 (1973), 97–112.
- [46] I. Nakamura, Inoue-Hirzebruch surfaces and a duality of hyperbolic unimodular singularities, I, *Math. Ann.*, 252 (1980), 221–225.
- [47] S. Nakano, On the inverse monoidal transformation, *Publ. Res. Inst. Kyoto Univ.*, 6 (1971), 483–502.
- [48] M. Namba, Moduli of open holomorphic maps of compact complex manifolds, *Math. Ann.*, 220 (1976), 65–76.
- [49] A. Newlander and L. Nirenberg, Complex analytic coordinates in almost complex manifolds, *Ann. Math.*, (2) 65 (1957), 391–404.
- [50] J.-P. Serre, Un théorème de dualité, *Comment. Math. Helv.*, 29 (1955), 9–26.
- [51] J.-P. Serre, Géométrie algébrique et géométrie analytique, *Ann. Inst. Fourier*, 6 (1955–1956), 1–42.
- [52] C. L. Siegel, Meromorphe Funktionen auf kompakten analytischen Mannigfaltigkeiten, *Nachr. Akad. Wiss. Göttingen*, 1955, 71–77.
- [53] T. Ueda, Compactification of $C \times C^*$ and $(C)^{*2}$, *Tohoku Math. J.*, 31 (1979), 81–90.
- [54] K. Ueno, Classification theory of algebraic varieties and compact complex spaces, *Lecture notes in math.* 439, Springer, 1975.

- [55] K. Ueno, Classification of algebraic varieties II. Algebraic threefolds of parabolic type, Int. Symp. Algebraic Geometry, Kyoto, 1977, 693–708.
- [56] K. Ueno, On algebraic fiber spaces of Abelian varieties, Math. Ann., 237 (1978), 1–22.
- [57] E. Viehweg, Canonical divisors and the additivity of the Kodaira dimension for morphisms of relative dimension one, Compositio Math., 35 (1977), 197–223.
- [58] E. Viehweg, Klassifikationstheorie algebraischer Varietäten der Dimension Drei, Compositio Math., 41 (1980), 361–400.
- [59] K. Yano, Differential geometry on complex and almost complex spaces, Macmillan, 1965.

73 (V.15) Complex Multiplication

A. Classical Theory

If the ratio ω_1/ω_2 of two periods ω_1, ω_2 of an \dagger elliptic function f belongs to an imaginary \dagger quadratic field K , then there exists an algebraic relation between $f(z)$ and $f(\lambda z)$ for any λ in K , and such an f is said to have **complex multiplication**. This phenomenon for the \dagger sn function with modulus $\sqrt{-1}$ was discovered by C. F. Gauss and was applied to the problem of dividing a \dagger lemniscate into five arcs of equal length. More generally, N. H. Abel showed that the *special dividing equation* of an sn function with complex multiplication is algebraically solvable. From a number-theoretic point of view, L. Kronecker conjectured that every \dagger Abelian extension of an imaginary quadratic number field K is determined by a *transform equation* of an elliptic function with complex multiplication by a number of K (1880). This is an analog of the fact, announced by Kronecker and proved by H. Weber, that every Abelian extension of the rational number field is a subfield of a \dagger cyclotomic field. Kronecker's work was continued by Weber [2], and his conjecture was proved by T. Takagi (1903) for $K = \mathbf{Q}(\sqrt{-1})$, by T. Takenouchi (1916) for $K = \mathbf{Q}(e^{2\pi i/3})$, and by Takagi (1920) for the general case using \dagger class field theory. H. Hasse [5] and M. Sugawara simplified the theory of complex multiplication, and Hasse noticed a relationship between complex multiplication and \dagger congruence zeta functions. Working from Hasse's idea, M. Deuring constructed the theory of complex multiplication purely algebraically

and determined \dagger Hasse's zeta function of an elliptic curve with complex multiplication.

For the rest of this article, K always stands for an imaginary quadratic field. Let L be a \dagger lattice group on the complex plane \mathbf{C} generated by ω_1, ω_2 , and let z be a complex variable. Define functions \wp, g_2, g_3 as follows: $\wp(z, L) = \wp(z; \omega_1, \omega_2) = z^{-2} + \sum ((z - \omega)^{-2} - \omega^{-2})$, $g_2(L) = g_2(\omega_1, \omega_2) = 60 \sum \omega^{-4}$, $g_3(L) = g_3(\omega_1, \omega_2) = 140 \sum \omega^{-6}$, where the sum \sum is over the elements ω of L except 0. Let \wp' be the derivative of \wp ; then $z \rightarrow (1, \wp(z), \wp'(z))$ is a one-to-one correspondence between the points on the complex torus \mathbf{C}/L and those on the \dagger elliptic curve E ; $X_0 X_2^2 = 4X_1^3 - g_2 X_1^2 X_0 - g_3 X_0^3$ in the projective plane. If the quotient ω_1/ω_2 generates K , then the ring of analytic endomorphisms of \mathbf{C}/L (i.e., the ring of endomorphisms of E) is isomorphic to a subring of the \dagger principal order \mathfrak{o} of K . In particular, if the lattice group L is an \dagger ideal of K (for $K \subset \mathbf{C}$; \rightarrow 347 Quadratic Fields), then the ring of endomorphisms coincides with \mathfrak{o} . The function $J(\tau) = J(E) = J(L) = 2^6 3^3 g_2(L)^3 / \Delta(L)$ ($\Delta(L) = g_2(L)^3 - 27g_3(L)^2$) of $\tau = \omega_1/\omega_2$, $\text{Im } \tau > 0$, is a \dagger modular function of level 1, and $J(E)$ is called the **invariant** of the elliptic curve E . If E has a complex multiplication, then $J(E)$ is an algebraic integer. Then the three main theorems of the classical theory of complex multiplication can be stated.

Theorem 1. Let h be the \dagger class number of K , and let $\mathfrak{a}_1, \dots, \mathfrak{a}_h$ be a set of representatives of ideal classes of K . Then $J(\mathfrak{a}_1), \dots, J(\mathfrak{a}_h)$ are exactly the conjugates of $J(\mathfrak{a}_1)$ over K , and $K(J(\mathfrak{a}_1))$ is the maximal unramified Abelian extension (the \dagger absolute class field) of K (\rightarrow 59 Class Field Theory).

Next we define the function f by

$$\begin{aligned} f(z; L) &= g_2 g_3 \Delta^{-1} \cdot \wp(z; L), \\ K &\neq \mathbf{Q}(\sqrt{-1}), \quad \mathbf{Q}(e^{2\pi i/3}), \\ &= g_2^2 \Delta^{-1} \cdot \wp(z; L)^2, \quad K = \mathbf{Q}(\sqrt{-1}), \\ &= g_3 \Delta^{-1} \cdot \wp(z; L)^3, \quad K = \mathbf{Q}(e^{2\pi i/3}). \end{aligned}$$

Theorem 2. Let \mathfrak{o} be the \dagger principal order of K , \mathfrak{m} be an integral ideal of K , and \mathfrak{a} be an arbitrary ideal of K . Choose a number ξ of K such that $\mathfrak{m} = \{\lambda \in \mathfrak{o} \mid \lambda \xi \in \mathfrak{a}\}$. Then $K(J(\mathfrak{a}), f(\xi; \mathfrak{a}))$ is the \dagger class field for the \dagger ray modulo \mathfrak{m} .

The number ξ in this theorem is obtained by $\mathfrak{a}^{-1} \mathfrak{m}(\xi) = \mathfrak{b}$, where \mathfrak{b} is an integral ideal belonging to the ideal class of $\mathfrak{a}^{-1} \mathfrak{m}$. Thus $f(\xi; \mathfrak{a})$ is a "special value" of an elliptic function as well as of a modular function.

The \dagger general law of reciprocity, the \dagger principal ideal theorem, and the ramification in the class field in theorem 2 can also be described

in terms of elliptic functions or elliptic curves. In general, the ring of endomorphisms of an elliptic curve defined over a field k of characteristic 0 is either \mathbf{Z} or an order of an imaginary quadratic field. On the other hand, if the characteristic of k is not zero, then the ring of endomorphisms may be an order in a definite \dagger quaternion algebra.

B. Complex Multiplication of an Abelian Variety

Following Kronecker's idea, D. Hilbert posed in his lecture at Paris (1900) the so-called 12th problem: to find an analytic function whose special values generate Abelian extensions over a given algebraic number field (\rightarrow 196 Hilbert). E. Hecke constructed unramified Abelian extensions of an imaginary biquadratic field using \dagger Hilbert modular functions [10]. After this, there was no notable development concerning the problem until the theory of complex multiplication was generalized to the case of \dagger Abelian varieties, which was made possible by progress in algebraic geometry, in particular A. Weil's geometric theory of Abelian varieties (G. Shimura and Y. Taniyama [11]). The following results have been obtained: Consider a triple (A, \mathfrak{X}, t) , consisting of an Abelian variety A defined over \mathbf{C} , a \dagger polarization \mathfrak{X} of A , and a point t of A . Call two such triples (A, \mathfrak{X}, t) and (A', \mathfrak{X}', t') isomorphic if an isomorphism of A onto A' maps \mathfrak{X} onto \mathfrak{X}' and t onto t' . Then there exists one and only one subfield k_0 of \mathbf{C} with the following property: In order for (A, \mathfrak{X}, t) and $(A^\sigma, \mathfrak{X}^\sigma, t^\sigma)$ to be isomorphic for an automorphism σ of \mathbf{C} , it is necessary and sufficient that σ fix all elements of k_0 . We call k_0 the **field of moduli** of (A, \mathfrak{X}, t) . If A is an elliptic curve E and if $t=0$, then $k_0 = \mathbf{Q}(J(E))$. In the higher-dimensional case, the field of moduli is generated by special values of a \dagger Siegel modular function.

If F is a totally imaginary number field that is a quadratic extension of a totally real field F_0 of degree n , then there exists a set $\{\varphi_1, \dots, \varphi_n\}$ of n different isomorphisms of F into \mathbf{C} such that $\bar{\varphi}_i = \varphi_j$ never occurs for $i \neq j$, where the bar denotes complex conjugation. We call $(F; \{\varphi_\lambda\})$ a **CM-type**.

If \mathfrak{a} is an ideal of F , then $L = \{(\varphi_1(\alpha), \dots, \varphi_n(\alpha)) \in \mathbf{C}^n \mid \alpha \in \mathfrak{a}\}$ is a lattice group in the n -dimensional complex linear space \mathbf{C}^n , and \mathbf{C}^n/L is analytically isomorphic to an Abelian variety A of dimension n in a complex projective space. The \dagger endomorphism ring $\mathfrak{A}(A)$ of A contains a ring that is isomorphic to the principal order \mathfrak{o} of F . Conversely, every

Abelian variety of dimension n such that $\mathfrak{A}(A)$ contains a ring isomorphic to \mathfrak{o} can be constructed in this way.

Let M be a normal extension of \mathbf{Q} containing F . Denote the Galois group of M by G and the subgroup of G corresponding to F by H , and put $S = \bigcup_\lambda \varphi_\lambda H$, where φ_λ stands for a prolongation of the previous φ_λ to G . Then the following three conditions are equivalent: (i) $\mathfrak{A}(A) \cong \mathfrak{o}$; (ii) A is simple; (iii) $H = \{\gamma \in G \mid S\gamma = S\}$. If $H^* = \{\delta \in G \mid \delta S = S\}$, then we can choose $\psi_\mu \in G$ such that $S = \bigcup_\mu H^* \psi_\mu$. If F^* is the subfield of M corresponding to H^* , then $(F^*; \{\psi_\mu\})$ is also a CM-type, and $F^* = \mathbf{Q}(\sum_\lambda \varphi_\lambda(\alpha) \mid \alpha \in F)$. Moreover, for an ideal \mathfrak{r} of F^* , $\prod_\mu \psi_\mu(\mathfrak{r})$ is an ideal of F .

Now let H_m be the group of ideals \mathfrak{r} of F^* that are relatively prime to the norm $N(\mathfrak{m})$ of an integral ideal \mathfrak{m} of F^* and for which there exists a number ξ of F with

$$\prod_\mu \psi_\mu(\mathfrak{r}) = (\xi), \quad N(\mathfrak{r}) = |\xi|^2, \quad \xi \equiv 1 \pmod{\mathfrak{m}}.$$

Then H_m is an \dagger ideal group modulo \mathfrak{m} in F^* .

Theorem 3. Assume that $\mathfrak{A}(A) \cong \mathfrak{o}$, and let \mathfrak{X} be an arbitrary polarization of A . Denote a point of A with $\mathfrak{m} = \{\lambda \in \mathfrak{o} \mid \lambda t = 0\}$ by t , and let k_m be the field of moduli of (A, \mathfrak{X}, t) . Then $k_m F^*$ is the \dagger class field over F^* corresponding to H_m .

The point t in theorem 3 always exists. In particular, if $\mathfrak{m} = \mathfrak{o}$, then $t=0$. When A is an elliptic curve E , F is an imaginary quadratic field and we have $F^* = K$, and theorem 3 coincides essentially with the content of theorems 1 and 2. If $n > 1$, $F = F^*$ holds only in special cases.

The theory of complex multiplication of A is closely related to the \dagger Hasse zeta function of A (\rightarrow 450 Zeta Functions).

References

- [1] L. Kronecker, Zur Theorie der elliptischen Funktionen, Monathber. Königlich Preuss. Akad. Wiss. Berlin, (1881), 1165–1172. (Werke IV, Teubner, 1929, 309–318; Chelsea, 1969.)
- [2] H. Weber, Lehrbuch der Algebra III, Vieweg, second edition, 1908 (Chelsea, 1961).
- [3] T. Takagi, Über eine Theorie des relativ-Abelschen Zahlkörpers, J. Coll. Sci. Imp. Univ. Tokyo, 41 (1920), 1–132.
- [4] R. Fueter, Vorlesungen über die singulären Moduln und die komplexe Multiplikation der elliptischen Funktionen, Teubner, I, 1924; II, 1927.
- [5] H. Hasse, Neue Begründung der komplexen Multiplikation, J. Reine Angew. Math., I, 157 (1926), 115–139; II, 165 (1931), 64–88.

- [6] R. Fricke, Lehrbuch der Algebra III, Vieweg, 1928.
- [7] M. Deuring, Die Typen der Multiplikatorringe elliptischer Funktionenkörper, Abh. Math. Sem. Univ. Hamburg, 14 (1941), 197–272.
- [8] M. Deuring, Die Struktur der elliptischen Funktionenkörper und die Klassenkörper der imaginären quadratischen Zahlkörper, Math. Ann., 124 (1952), 393–426.
- [9] M. Deuring, Die Klassenkörper der komplexen Multiplikation, Enzykl. Math., Bd. I₂, Heft 10, Teil II, 23, 1958.
- [10] E. Hecke, Höhere Modulformen und ihre anwendung auf die Zahlentheorie, Math. Ann., 71 (1912), 1–37; Über die Konstruktion relativ-Abelscher Zahlkörper durch Modulformen von zwei Variablen, Math. Ann., 74 (1913), 465–510. (Mathematische Werke, Vandenhoeck & Ruprecht, 1959, 21–58, 69–114.)
- [11] G. Shimura and Y. Taniyama, Complex multiplication of Abelian varieties and its applications to number theory, Publ. Math. Soc. Japan, 1961.
- [12] G. Shimura, Construction of class fields and zeta functions of algebraic curves, Ann. Math., (2) 85 (1967), 58–159.
- [13] A. Borel et al., Seminar on complex multiplication, Lecture notes in math. 21, Springer, 1966.
- [14] G. Shimura, Introduction to the arithmetic theory of automorphic functions, Publ. Math. Soc. Japan and Princeton Univ. Press, 1971.

74 (II.11) Complex Numbers

A. Algebraic Properties of Complex Numbers

A **complex number** is an expression of the form $a + ib$ with arbitrary real numbers a and b and the **imaginary unit** i . Writing $\alpha = a + ib$, $\beta = c + id$, we define $\alpha = \beta$ if and only if $a = c$ and $b = d$. As regards algebraic operations with complex numbers, we define $\alpha + \beta = (a + c) + i(b + d)$, $\alpha - \beta = (a - c) + i(b - d)$, $\alpha\beta = (ac - bd) + i(ad + bc)$, and for $\beta \neq 0$, i.e., $c^2 + d^2 \neq 0$, $\alpha/\beta = (ac + bd)/(c^2 + d^2) + i(bc - ad)/(c^2 + d^2)$. Then the addition and multiplication thus defined obey commutative, associative, and distributive laws, and complex numbers form a \dagger commutative field with $0 = 0 + i0$ and $1 = 1 + i0$ as its zero element of addition and identity element of multiplication, respectively. The set of all complex numbers is usually denoted by \mathbf{C} .

By assigning to each real number a a complex number $a + i0$, algebraic operations on real numbers are carried into those of the corresponding complex numbers. That is to say, the field \mathbf{R} of all real numbers is mapped isomorphically into the field \mathbf{C} of all complex numbers. By identifying a with $a + i0$, we are taking \mathbf{R} as a subfield of \mathbf{C} . Also, $0 + i1$ will be denoted simply by i . From the previous definition of algebraic operations, it follows that $i^2 = -1$. Furthermore, since $\alpha = a + ib = (a + i0) + (b + i0)(0 + i1)$, $a + ib$ is not a mere symbolic expression but can also be regarded as the outcome of algebraic operations on a, b, i in \mathbf{C} . The **real** and **imaginary parts** of a complex number $\alpha = a + ib$ are, by definition, a and b , denoted by $\operatorname{Re} \alpha$ and $\operatorname{Im} \alpha$, respectively. A complex number that is not a real number is sometimes called an **imaginary number**; in particular, a complex number α with $\operatorname{Re} \alpha = 0$ is called a **purely imaginary number**. For each complex number $\alpha = a + ib$, we define its **conjugate complex number** as $a - ib$, and denote it by $\bar{\alpha}$. We then have $\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta}$ and $\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}$. The mapping $\alpha \rightarrow \bar{\alpha}$ is an \dagger automorphism of \mathbf{C} which leaves each element of \mathbf{R} invariant. Also, the following relations hold: $\operatorname{Re} \alpha = (\alpha + \bar{\alpha})/2$ and $\operatorname{Im} \alpha = (\alpha - \bar{\alpha})/(2i)$.

Regarded as an overfield of \mathbf{R} , \mathbf{C} is an \dagger extension field of \mathbf{R} of degree 2 obtained by the adjunction of i , which is a root of an irreducible equation $x^2 + 1 = 0$. The important algebraic property of \mathbf{C} is that it is \dagger algebraically closed. Namely, for any polynomial $f(x)$ with coefficients in \mathbf{C} , the equation $f(x) = 0$ possesses at least one root in \mathbf{C} (\dagger Gauss's fundamental theorem of algebra).

B. Topology of \mathbf{C}

The **absolute value** (or **modulus**) of a complex number $\alpha = a + ib$, denoted by $|\alpha|$, is by definition $|\alpha| = \sqrt{a^2 + b^2} = \sqrt{\alpha\bar{\alpha}}$. If α is real, then the absolute value of α in the sense of complex numbers is identical to the one in the sense of real numbers. It always holds that $|\alpha| \geq 0$ and $|\alpha| = 0 \Leftrightarrow \alpha = 0$. It follows further that $|\alpha + \beta| \leq |\alpha| + |\beta|$, $|\alpha\beta| = |\alpha||\beta|$, and $|\alpha| = |\bar{\alpha}|$.

For each pair of complex numbers α and β , define $\rho(\alpha, \beta) = |\alpha - \beta|$. Then with $\rho(\alpha, \beta)$ as the \dagger distance function, \mathbf{C} satisfies the axioms for a \dagger metric space, and in particular, $\lim_{n \rightarrow \infty} \alpha_n = \alpha_0 \Leftrightarrow \lim \rho(\alpha_n, \alpha_0) = 0 \Leftrightarrow \lim |\alpha_n - \alpha_0| = 0 \Leftrightarrow (\lim \alpha_n = a_0 \text{ and } \lim \beta_n = b_0) \text{ (where } \alpha_n = a_n + ib_n, a_0 = a_0 + ib_0)$. From this it is easily seen, as in the case of the set \mathbf{R} of all real numbers, that \mathbf{C} also becomes a \dagger locally compact and \dagger complete metric space.

With respect to this topology, the four oper-

ations (except for division by zero) are continuous: $\alpha_n \rightarrow \alpha_0$ and $\beta_n \rightarrow \beta_0$ imply $\alpha_n + \beta_n \rightarrow \alpha_0 + \beta_0$; $\alpha_n - \beta_n \rightarrow \alpha_0 - \beta_0$; $\alpha_n \beta_n \rightarrow \alpha_0 \beta_0$; and $\alpha_n / \beta_n \rightarrow \alpha_0 / \beta_0$ (where in the last case we assume $\beta_n \neq 0$ and $\beta_0 \neq 0$). Thus \mathbb{C} becomes a topological field. Furthermore, the assignment $\alpha \rightarrow \bar{\alpha}$ gives a continuous mapping $\mathbb{C} \rightarrow \mathbb{C}$, a homeomorphic automorphism of \mathbb{C} .

C. The Complex Plane

If in a plane to which is assigned rectangular coordinate axes a complex number $\alpha = a + ib$ is represented by a point (a, b) , then the plane is called the **complex (number) plane** (**Gauss-Argand plane** or **Gaussian plane**) (Fig. 1), and the point representing α is called simply the point α . The abscissa and ordinate axes are called the **real** and **imaginary axes**, respectively. A point $\alpha = a + ib$ can be represented by † polar coordinates r, θ with the origin and the real axis as the pole and the generating line, respectively, where $r = \sqrt{a^2 + b^2}$ is the absolute value $|\alpha|$ of α and θ is the **argument** (or **amplitude**), denoted by $\arg \alpha$, of α . The argument of α is uniquely determined mod 2π if $a \neq 0$ and is an arbitrary real number if $\alpha = 0$.

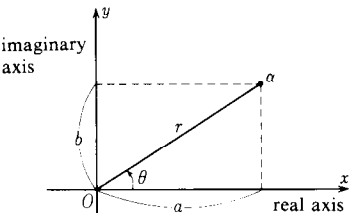


Fig. 1

The absolute value $|\alpha|$ of a complex number α , regarded as a vector from the origin to the point α , is the length of this vector. For complex numbers α and β , to the sum of the vectors α and β corresponds the sum $\alpha + \beta$ of the complex numbers. A complex number α , in terms of its absolute value r and argument θ , is expressed as $\alpha = r(\cos \theta + i \sin \theta)$, which is called the **polar form** of α . For polar forms the following hold: $\bar{\alpha} = r(\cos \theta - i \sin \theta) = r(\cos(-\theta) + i \sin(-\theta))$; $\alpha^{-1} = \bar{\alpha} / |\alpha|^2 = r^{-1}(\cos(-\theta) + i \sin(-\theta))$ ($\alpha \neq 0$); and $\alpha_1 \alpha_2 = r_1 r_2 (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2))$, where $|\alpha_j| = r_j$ and $\arg \alpha_j = \theta_j$ for $j = 1, 2$. This last relation when $r_1 = r_2 = 1$ is called **De Moivre's formula**. The n th roots of unity in \mathbb{C} are given by $\rho_j = \cos 2\pi j/n + i \sin 2\pi j/n$ ($j = 0, 1, \dots, n-1$) (Fig. 2).

In the complex plane, the mapping $\alpha \rightarrow \bar{\alpha}$ corresponds to the † reflection of the plane in the real axis, $\alpha \rightarrow \alpha + \beta$ to the parallel † translation along a vector β , $\alpha \rightarrow \alpha\beta$ ($\beta \neq 0$) to the † rotation through the angle $\arg \beta$ followed by

a † homothetic transformation with center 0 and constant ratio $|\beta|$, and $\alpha \rightarrow \bar{\alpha}^{-1}$ to the † inversion with respect to the **unit circle** $\{\alpha \mid |\alpha| = 1\}$.

The distance $\rho(\alpha, \beta) = |\alpha - \beta|$ between α and β in \mathbb{C} coincides with their Euclidean distance, provided that α and β are regarded as points in the Euclidean plane, so that the complex plane is † isometric, and accordingly homeomorphic, to the Euclidean plane.

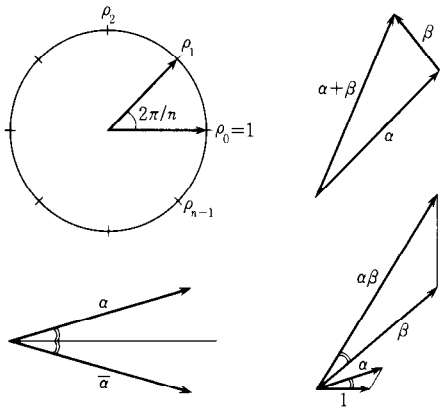


Fig. 2

D. The Complex Sphere

In the rest of this article, P denotes the complex plane and Σ denotes the sphere of radius 1, with 0, the origin of P , as its center. The points $N(0, 0, 1)$ and $S(0, 0, -1)$ of Σ will be called the **north** and **south pole**, respectively (Fig. 3), where the 1st and 2nd coordinate axes are the real and imaginary axes of P , respectively, and the 3rd coordinate axis is orthogonal to P . A straight line from N through a point z (a complex number) in P intersects Σ at a point $Z = (x_1, x_2, x_3)$ different from N , where $z = (x_1 + ix_2)/(1 - x_3)$, $x_1 = (z + \bar{z})/(1 + |z|^2)$, $x_2 = (z - \bar{z})/(i(1 + |z|^2))$, and $x_3 = (|z|^2 - 1)/(|z|^2 + 1)$. The mapping $z \rightarrow Z$ is called a **stereographic projection** from N , by means of which P and $\Sigma - \{N\}$ become † conformally equivalent to each other. Consequently z can be represented by a point Z of $\Sigma - \{N\}$, and Σ thus used is called a **complex sphere** (or **Riemann sphere**). Let us adjoin to the complex plane P a new element, denoted by ∞ , called the **point at infinity**, which corresponds to the only exceptional point N of Σ . The topology of the complex plane with ∞ can be introduced by the corresponding topology of the Riemann sphere. Indeed, the family of all the sets $\{z \mid |z| > M\} \cup \{\infty\}$ for $M > 0$ forms a † local base around ∞ . By introducing local complex coordinates $\zeta = 1/z$ into the neighborhoods of ∞ , each element of this local base is represented as $\{\zeta \mid |\zeta| < M^{-1}\}$, in which the

convention $\zeta = 0$ is adopted for $z = \infty$. The complex sphere thus defined can be regarded as a \dagger Riemann surface (i.e., a 1-dimensional \dagger complex manifold).

The complex plane (complex sphere) whose points are represented by a variable z or w is called a **z-plane** or a **w-plane** (a **z-sphere** or a **w-sphere**).

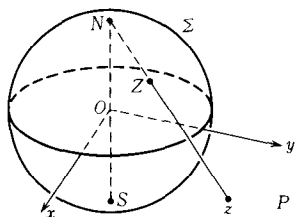


Fig. 3

E. Linear Fractional Functions

Given complex numbers a, b, c , and d with $ad - bc \neq 0$, we define a **linear fractional function** (or simply **linear function**) by

$$w = \frac{az + b}{cz + d}. \quad (1)$$

As a mapping from the z -sphere into the w -sphere, this linear function is called a **Möbius transformation** (**linear fractional** or simply **linear transformation**). The usual linear transformation, i.e., the one with $c = 0$ in the present case, is sometimes distinguished as an **entire linear transformation**. Since (1) depends only on the proportion $a:b:c:d$, we can assume $ad - bc = 1$ without loss of generality.

The transformation (1) is \dagger holomorphic and \dagger univalent on the whole z -sphere with only one exceptional \dagger pole at $-d/c$ (∞ if $c = 0$) of order 1, and the inverse of (1) is also a linear fractional function. The set of all linear transformations forms a \dagger group with composition of transformations as the group operation. One of its subgroups is the \dagger modular group.

Linear transformations carry any circle of the complex plane (or of the Riemann sphere) into a circle of the same plane (or of the same sphere) if we adopt the convention that straight lines are a special kind of circle. (In the case of a Riemann sphere, no such convention is necessary.) Given on a plane a circle with center o and radius r and two points p and p' on a half-line issuing from o satisfying $op \cdot op' = r^2$, the points p and p' are called **symmetric points** (or **reflection points**) with respect to the circle. The transformation $p \rightarrow p'$ is called the **inversion** with respect to this circle. In the complex plane, let z and z' be symmetric points with respect to a circle C . Suppose that by a linear transformation, z, z' and

C are carried to points w, w' and a circle D , respectively; then w and w' become symmetric with respect to the circle D (**principle of reflection**). Thus symmetricity is invariant under linear transformations. Also the \dagger anharmonic ratio of any four points z_1, z_2, z_3 , and z_4 , $(z_1, z_2; z_3, z_4) = (z_1 - z_3)/(z_1 - z_4) : (z_2 - z_3)/(z_2 - z_4)$, is invariant under a linear transformation; i.e., $(z_1, z_2; z_3, z_4) = (w_1, w_2; w_3, w_4)$ holds, where w_j is the image of z_j under a linear transformation ($j = 1, 2, 3, 4$).

F. Normal Forms of Linear Transformations

There exist fixed points of the transformation (1) on the Riemann sphere, i.e., points satisfying $z = (az + b)/(cz + d)$. The number of fixed points is 2 or 1, except when $w = z$. If the transformation has two fixed points, they will be denoted here by p and q . The natural convention $p = q$ is adopted if the transformation has one fixed point. If $c = 0$, then p or q is ∞ , and furthermore if $c = a - d = 0$, then p and q are both ∞ .

For unequal finite p and q , (1) can be rewritten in the following normal form:

$$\frac{w-p}{w-q} = \alpha \frac{z-p}{z-q}, \quad \alpha = \frac{a-cp}{a-cq} \neq 1,$$

in which, according as $\arg \alpha = 0$, $|\alpha| = 1$, or otherwise, (1) is called a **hyperbolic** (Fig. 4), **elliptic** (Fig. 5), or **loxodromic transformation**, respectively. This classification can be applied also for finite p and infinite q , i.e., to the transformation $w - p = \alpha(z - p)$. Furthermore, for $p = q \neq \infty$, (1) is rewritten in the following form:

$$\frac{1}{w-p} = \frac{1}{z-p} + \beta, \quad \beta = \frac{c}{a-cp}.$$

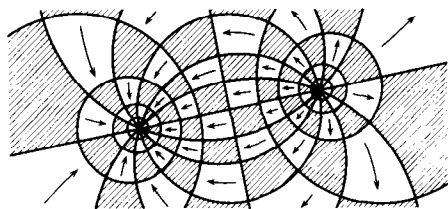


Fig. 4
Hyperbolic transformation.

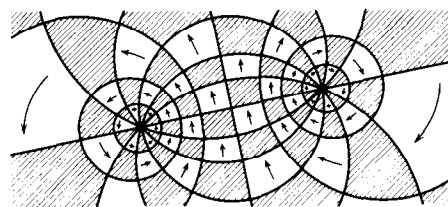


Fig. 5
Elliptic transformation.

In this case (1) is called a **parabolic transformation** (Fig. 6). For $p = q = \infty$, i.e., if $w = z + \beta$, (1) is also called parabolic. We can easily determine to which class (1) belongs from the discriminant $D = (a + d)^2 - 4$ of the quadratic equation $cz^2 - (a - d)z - b = 0$ obtained from $z = (az + b)/(cz + d)$ with $ad - bc = 1$ by multiplying both sides by $cz + d$. If $a + d$ is real, then according as $D > 0$, < 0 , or $= 0$, (1) is hyperbolic, elliptic, or parabolic, respectively, and if $a + d$ is not real, then the transformation is loxodromic.

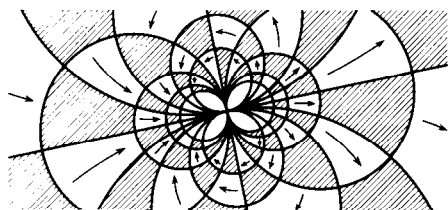


Fig. 6
Parabolic transformation.

Let D and D' be two arbitrary circular disks. Then there always exists a linear transformation which gives a one-to-one conformal mapping from D onto D' . Conversely, any mapping with this property is given only by linear transformations (provided that the half-plane having a straight line together with the point at infinity as its boundary is regarded as a closed disk), which are uniquely determined by giving three points a, b, c from the boundary of D and as their corresponding points, three arbitrary points a', b', c' from the boundary of D' .

G. The Poincaré Metric

Since conformal mappings from the domain $|z| < 1$ onto $|w| < 1$ are given by the transformations $w = \varepsilon(z - z_0)/(1 - \bar{z}_0 z)$ ($|\varepsilon| = 1$, $|z_0| < 1$) (\rightarrow Appendix A, Table 13), for corresponding z and w it holds that

$$\frac{|dw|}{1 - |w|^2} = \frac{|dz|}{1 - |z|^2}. \quad (2)$$

$|dz|/(1 - |z|^2)$ is called **Poincaré's differential invariant**. With a metric having $ds = |dz|/(1 - |z|^2)$ as its line element, the unit disk $|z| < 1$ becomes a non-Euclidean space in the sense of Lobachevskii, and the metric is called the **Poincaré metric**. Furthermore, since the transformations (2) leave the length of curves invariant, they can be regarded as motions in this space, where the geodesic through two points z_1 and z_2 is the circular arc orthogonal to the unit circle. If we denote the intersections of the arc with the unit circle by z_3 and z_4 , then the non-Euclidean distance between two points

z_1 and z_2 along the geodesic is given by $(1/2)\log(z_1, z_2; z_3, z_4)$, provided that the points z_4, z_1, z_2, z_3 are arranged on the arc in this order (\rightarrow 285 Non-Euclidean Geometry).

References

- [1] C. F. Gauss, Werke I–XII, Göttingen, 1863–1931.
- [2] N. Bourbaki, *Eléments de mathématique* III. Topologie générale, ch. 8. Nombres complexes, Actualités Sci. Ind., 1235b, Hermann, third edition, 1963; English translation, General topology, pt. 2, Addison-Wesley, 1966.
- [3] L. V. Ahlfors, *Complex analysis*, McGraw-Hill, second edition, 1966.
- [4] C. Carathéodory, *Funktionentheorie I*, Birkhäuser, 1950; English translation, *Theory of functions of a complex variable I*, Chelsea, 1954.

75 (XVI.2) Computers

A. History

Since the beginning of civilization, people have utilized tools for aiding computation. In Japan, bamboo computing rods were used in the 7th century; before the close of the 16th century, the **abacus** was imported from China (\rightarrow 230 Japanese Mathematics (Wasan)). The first **calculator** capable of performing the four arithmetic operations automatically was designed by W. Schickard of Tübingen University (1623). After that, B. Pascal independently made his famous adding machine; this was improved by G. W. Leibniz so that it was able to execute multiplications and divisions.

Most calculators manufactured today are electronic; however, many of these machines are not "programmable" and hence require at every step of computation a manual operation for specifying the machine operations to be executed. On the other hand, modern automatic computers can execute automatically a sequence of operations according to a given program without any manual intervention.

The automatic computer was conceived by the English mathematician C. Babbage in the 19th century, but mechanical engineering at that time was not advanced enough to allow the construction of such a computer. His idea was first realized by the relay computers Z3 of K. Zuse (1941) and Mark I of Harvard University (1944). In 1947, the first **electronic computer**, ENIAC, appeared, in which vacuum

tubes were utilized instead of mechanical components. Since then, the capabilities of computers have increased rapidly and it can be said that we now live in the "computer age."

B. Principles of Modern Computers

Information processing in a computer is based on communication among its constituents by electric signals. In **digital computers**, information is encoded as a sequence of binary numbers 0, 1, whereas continuous values are allowed in **analog computers**. The minimum quantity of information in digital computers is therefore a binary digit, called a **bit**. Since continuous values fluctuate on account of electric noise in the circuits, digital computers have the advantage of maintaining high precision during computation.

The binary numbers 0, 1 are represented in practice by two distinct electric signals: two distinct voltages, two distinct phases of alternating current, the existence or nonexistence of a pulse, etc. A system of circuits is called **synchronous** when it contains a clock, a generator of periodic pulses, which synchronizes the transmission and transformation of information. In an **asynchronous** system, circuits execute each step of information processing independently and advance to the next step after verifying the termination of the preceding step.

In ENIAC, a program was inserted into the computer by carefully connecting many control lines on plugboards. Since this caused much trouble in inserting the different programs needed for different jobs, J. von Neumann proposed encoding programs in sequences of binary digits and storing these programs in the memory unit of the computer. His principle, the **stored program principle**, was realized by EDSAC (1949) and has since been widely used.

An automatic computer in general consists of the following five units: the arithmetic unit, memory, control, input, and output (Fig. 1). These units are interconnected by wires that exchange information in the course of computation.

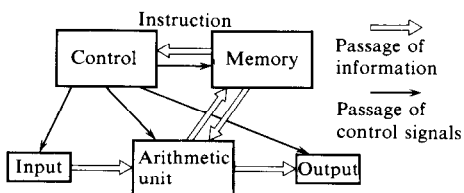


Fig. 1
Construction of electronic computers.

The **arithmetic unit** consists mainly of several memory registers and operational circuits associated with them. Each register stores a binary number of n bits (usually $16 \leq n \leq 64$). An important building block of the operational circuit is the basic adder of 1-bit numbers. A parallel adder of n -bit numbers can be obtained by connecting n copies of the basic adder. Alternatively, a sequential adder can be composed of a single basic adder which is utilized repeatedly to sum up bit by bit two binary numbers from the lowest bit. Subtraction is usually carried out by adding the complement of the subtrahend. Multiplication is realized by shifting the multiplicand to the left and adding it to the intermediate sum; this addition may be omitted, depending on the relevant bit of the multiplier. In division, we shift the dividend to the left and if possible subtract the divisor from it. By recording at each step whether or not the subtraction is possible, we obtain the quotient.

The **memory unit** stores the instructions and given data as well as the necessary data obtained in the course of computation. It is divided into many registers, each of which contains a number and is referred to by means of a serial number called an **address**. At present, high-speed memory units are usually made up of magnetic cores or integrated circuits. As auxiliary large-scale memory, there are also magnetic drums, disks, and magnetic tapes, etc. Other elements under investigation are extreme-low-temperature elements, chemical elements, optical elements, etc.

The **control unit** repeats the following operations consecutively: (1) takes an instruction from the memory location indicated by the sequential control counter, (2) gets a data word from the memory according to the address part of the instruction, (3) decodes the function part of the instruction and sends control signals to appropriate circuits, and (4) increases the content of the sequential control counter by 1 and, after receiving *end* signals from the arithmetic unit, returns to step (1). For these purposes the control unit contains a counter, a decoder, an encoder to send control signals, and a register to store the instruction to be executed.

The arithmetic, control, and memory units form the **central processor** of the computer. In contrast to the central processor, the input and output units are called **peripheral devices**. An input device receives the necessary information (a program and data); a card reader reads punched cards, and a teletypewriter sends signals directly into the computer. An output device presents the results obtained by the computer. The results are usually printed by a teletypewriter or line printer. Magnetic tape

units can be considered as auxiliary input-output devices, since tape reels are removable. The tape units can accept prepared input data as well as record the results, the output.

These devices are operated according to given input-output instructions under the control of the central processor. However, since every input-output device contains mechanical components and is extremely slow in comparison with the central processor, a large-scale computer is often accompanied by satellite computers which undertake the control of input-output devices.

C. Instructions and Programming

In stored-program operation, a computer performs a sequence of calculations according to given instructions. There are various types of instructions, but those most frequently used are **single-address instructions**, each of which contains a single-address part designating an operand.

A **program** is a finite sequence of instructions arranged suitably for the required computation. **Programming**, or making a program, is therefore the task of decomposing the required computation into elementary steps each of which corresponds to an instruction.

Every instruction is represented in a computer by a number, a numeric code, which is determined in a definite way by the construction of the control unit. Before starting computation, instructions thus encoded are stored in the memory. In this sense, a program is a sequence of numbers. This sequence is called a **machine-language program**.

A program is usually divided into several blocks, called subprograms or **subroutines**. Some subroutines are made in advance, especially those for frequently required jobs such as evaluating elementary functions and manipulating input-output devices, etc. The system of these ready-made programs is called **software**, in contrast to the **hardware** (i.e., mechanico-electronic equipment) of the computer. Quite often basic routines are **microprogrammed**, i.e., written in a simple code and stored in a fast read-only memory (ROM). In this case the set of built-in programs is called **firmware**. The handiness of a computer depends mainly on close matching of the software and hardware.

Programs are usually written in certain forms called **external languages** which are easy to master. A problem-oriented language is an advanced external language in which ordinary arithmetic expressions are available with slight modifications. Programs written in these languages are translated into machine languages by program input routines. The translator for

a problem-oriented language is called the **compiler**.

A compiler accepts several macroinstructions. Moreover, it is equipped with the following facilities: (i) the ability to translate arithmetic expressions into machine language; (ii) the ability to generate linkages to various ready-made subroutines according to certain simple indications; (iii) the ability to automatically allocate programs, subroutines, and data in the memory; (iv) the ability to check automatically the syntactic correctness of programs. Thus it accepts an external form such as

```
if  $x \geq 0$  then printreal (SQRT( $x$ ))
```

```
    else printstring ('negative').
```

Another important translator is the **assembler**, which translates mnemonic codes of instructions (*add* for addition, etc.) into their numeric codes according to a given table. It allows us to utilize symbols for specifying addresses. It also converts decimal numbers into binary and generates certain segments of the program from rather simple indications. Compilers and assemblers are important constituents of software.

The large-scale high-speed computers that have recently appeared have made software systems inevitably more complex. There is now software, so-called monitors or **operating systems**, that supervises the uninterrupted processing of many programs. Some operating systems coordinate several assemblers and compilers so that several languages can be mixed in writing a program. Examples of other important software are mathematical (numerical) software and database management systems.

D. Mathematical Models of Computers

An operational circuit in the arithmetic unit is usually constructed from basic elements (logical gates), each of which performs a certain operation on the binary signals. Thus the construction of a circuit from basic elements is represented by a composition of a logical function defined over the set $\{0, 1\}$ from a given set of basic functions. Post [3] considered functional composition without feedback loops and established a general criterion for a given set of logical functions to be complete, i.e., to be capable of generating the whole set of logical functions. Strictly, however, a logical gate takes a definite time delay to perform its operation. Therefore Kudryavtsev [4] proposed as a model of a logical gate a pair (f, d) of a logical function f and a non-

negative integer d : f represents the operation of the gate and d represents its delay. He defined the feedback-free composition of such functions with delays and gave a completeness criterion for a set of logical functions with delays. His study has been extended in various directions by Loomis [5], Nozaki [6], and Rosenberg [7].

When a circuit contains memory elements or feedback loops, its function is suitably represented by an \dagger automaton. Hence the design and analysis of circuits with memory elements, state-minimization, and equivalence checking [8], decomposition into simpler components [9, 10], verification of completeness [11], etc., can be studied in terms of automata (\rightarrow 31 Automata). A computer itself can be considered to be a finite automaton, since it has a finite number of memory elements and its behavior is completely determined by the content of the memory (its internal state) and the inputs. However, since a modern computer has an enormous memory and contains replaceable parts, such as magnetic tapes, it is more adequately represented by an infinite model, such as a \dagger Turing machine.

Turing machines are capable of simulating many intellectual activities governed by formal rules. It is believed that any well-defined finite algorithm can be simulated by a Turing machine (Church's thesis.) Thus any computer can be simulated by a Turing machine, and hence it is unable to resolve those decision problems which are unsolvable for Turing machines. For instance, no program can decide in finite steps whether a given program written in a computer language, say FORTRAN or PASCAL, eventually stops or not. On the other hand, a modern computer can simulate a universal Turing machine provided that its memory can be extended unboundedly by supplying magnetic tapes. Hence any well-defined finite algorithm can be simulated by a computer unless the limitation of memory capacity hampers its accomplishment.

E. Mathematical Theory of Programming Languages

Along with the development of software, programming techniques have gradually accumulated. For instance, we now have an almost satisfactory method of translating arithmetic expressions into machine language. However, we still lack a general theory to cover effectively a wide array of programming problems. The design of an adequate metalanguage

is still an important problem, if the word "adequate" implies complete description of syntax and semantics of problem-oriented languages. An interesting problem is to define and suitably classify grammars with respect to their capabilities of forming languages. Such research is an important branch of **mathematical linguistics**. N. Chomsky [12] has investigated this problem, starting from research on natural languages, and has given formal definitions of the grammars (\rightarrow 31 Automata D).

\dagger Context-free grammars (type-2 languages) play an important role in the theory of software as well as in Chomsky's syntax, since they are powerful enough to describe the parenthesis structure and simple and easy to manipulate. Some variants of context-free grammar have been proposed for attaining high efficiency in the automatic appraisal of programs [13, 14] or for increasing the programs' generative power [15].

F. Branches of Information Science

Many fields related to computers now make up the **information sciences** or **informatiques**.

Important mathematical theories born or developed in the information sciences are:

(1) Design and analysis of hardware devices:

\dagger Boolean algebra [16], switching theory [17], and theory of \dagger automata [18, 19].

(2) Design and analysis of programming language; theory of formal languages [14, 21].

(3) Design and analysis of algorithms; \dagger numerical analysis, theory of \dagger complexity of computation, theory of \dagger data processing.

(4) Mathematical foundation of programming; logical verification of correctness and equivalence of programs [22], \dagger recursive function theory, \dagger decision problems.

Applications of the information sciences are found in diversified fields, such as statistics, operations research, mathematical psychology, econometrics, jurimetrics, and behaviorimetrics.

References

- [1] H. H. Goldstine, *The computer from Pascal to von Neumann*, Princeton University Press, 1972.
- [2] W. Eicken and E. H. Zuse, *Elektronische Datenverarbeitung*, 2 (1962), 74–75.
- [3] E. L. Post, *Introduction to a general theory of elementary propositions*, *Amer. J. Math.*, 43 (1921), 163–185.
- [4] V. B. Kudryavtsev, *A completeness theorem for a class of feedback-free automata* (in Russian), *Problemy Kibernet.*, 8 (1962), 91–115.

- [5] H. H. Loomis, Jr., Completeness of sets of delayed-logic devices, *IEEE Trans., EC-14* (1965), 157–172.
- [6] T. Hikita and A. Nozaki, A completeness criterion for spectra, *SIAM J. Comput.*, 6 (1977), 285–297.
- [7] I. G. Rosenberg, Completeness properties of multiple-valued logic algebras, *Computer Science and Multiple-Valued Logic, Theory and Applications*, D. C. Rine (ed.), North-Holland, 1977, 142–186.
- [8] S. Eilenberg, *Automata, languages and machines*, Academic Press, vol. A, 1974; vol. B, 1976.
- [9] J. Hartmanis and R. E. Stearns, *Algebraic structure theory of sequential machines*, Prentice-Hall, 1966.
- [10] A. Nozaki, Practical decomposition of automata, *Information and Control*, 36 (1978), 275–291.
- [11] M. A. Arbib, *Theories of abstract automata*, Prentice-Hall, 1969.
- [12] N. Chomsky, On certain formal properties of grammars, *Information and Control*, 2 (1959), 137–167.
- [13] J. E. Hopcroft and J. D. Ullman, *Formal languages and their relation to automata*, Addison-Wesley, 1969.
- [14] A. V. Aho and J. D. Ullman, *The theory of parsing, translation and compiling*, Prentice Hall, vol. 1, 1972; vol. 2, 1973.
- [15] A. Salomaa, *Formal languages*, Academic Press, 1973.
- [16] J. Kuntzmann, *Algebra de Boole*, Dunod, 1968.
- [17] S. Muroga, *Logic design and switching theory*, Wiley, 1979.
- [18] M. L. Minsky, *Computation: Finite and infinite machines*, Prentice-Hall, 1967.
- [19] J. E. Hopcroft and J. D. Ullman, *Introduction to automata theory, languages and computation*, Addison-Wesley, 1979.
- [20] Z. Manna, *Mathematical theory of computation*, McGraw-Hill, 1974.
- [21] W. A. Wulf, M. Shaw, P. N. Hilfinger, and L. Flon, *Fundamental structures of computer science*, Addison-Wesley, 1981.

76 (VI.17) Conformal Geometry

A. Möbius Geometry

We represent an n -dimensional sphere S^n as the \dagger quadric hypersurface $S^n: x_1^2 + x_2^2 + \dots + x_n^2 - 2x_0x_\infty = 0$ in an $(n+1)$ -dimensional real \dagger projective space \mathbf{P}^{n+1} , where the (x_α) are

\dagger homogeneous coordinates in \mathbf{P}^{n+1} . We denote by $M(n)$ the group of all \dagger projective transformations of \mathbf{P}^{n+1} that leave S^n invariant. Then the transformation group $M(n)$ acts on S^n . The pair $(S^n, M(n))$ is called the **conformal geometry** or **Möbius geometry**. We call S^n an n -dimensional **conformal space**, a transformation belonging to $M(n)$ a **Möbius transformation**, and $M(n)$ the **Möbius transformation group**. Every point of the projective space \mathbf{P}^{n+1} corresponds to a **hypersphere** on S^n . For example, if a point A lies outside of S^n , then the intersection of S^n and the \dagger polar hyperplane of A with respect to S^n is an $(n-1)$ -dimensional sphere S^{n-1} , and the point A corresponds to this **real hypersphere** S^{n-1} . Similarly, if a point A lies on S^n , it corresponds to a **point hypersphere**, and if a point A lies inside of S^n , then A corresponds to an **imaginary hypersphere**. We sometimes identify the point A with the corresponding hypersphere. For any two points $A=(a_\alpha)$ and $B=(b_\alpha)$, we put $AB=a_1b_1+a_2b_2+\dots+a_nb_n-(a_0b_\infty+a_\infty b_0)$ and call it the **inner product** of the two hyperspheres A and B . The **angle** θ between two intersecting real hyperspheres A and B is defined by $\cos \theta = AB/(\sqrt{A^2} \cdot \sqrt{B^2})$. This angle is invariant under the Möbius transformation.

In the projective space \mathbf{P}^{n+1} , we take a \dagger frame $(A_0, A_1, \dots, A_n, A_\infty)$ that satisfies the conditions

$$A_0^2 = A_0A_j = A_iA_\infty = A_\infty^2 = 0, \quad A_0A_\infty = -1,$$

$$A_iA_j = g_{ij}, \quad i, j = 1, 2, \dots, n,$$

where (g_{ij}) is a positive definite matrix. We see that A_0 and A_∞ are points on S^n and each A_i is a real hypersphere passing through these two points. Every hypersphere X of S^n can be written as a linear combination of the A_i : $X = u_0A_0 + u_1A_1 + \dots + u_nA_n + u_\infty A_\infty$. That is, X is represented by \dagger projective coordinates (u_α) with respect to the above frame. We call these homogeneous coordinates (u_α) $(n+2)$ -**hyperspherical coordinates** of the hypersphere X . If we use these coordinates, the inner product of two hyperspheres $X=(u_\alpha)$ and $Y=(v_\alpha)$ is given by $XY = \sum_{i,j=1}^n g_{ij}u_iv_j - (u_0v_\infty + u_\infty v_0)$.

The Möbius transformation group $M(n)$ is a topological group with two \dagger connected components. If we denote by $M_0(n)$ the maximal connected subgroup of $M(n)$ and by H the subgroup of $M(n)$ that leaves invariant a real hypersphere of S^n , then the set E of all real hyperspheres of S^n can be identified with the \dagger homogeneous space $M_0(n)/H$. The group H also consists of two connected components. If we denote by H_0 the maximal connected subgroup of H , the homogeneous space $\tilde{E} = M_0(n)/H_0$ is a two-fold \dagger covering space of E .

For each real hypersphere $A \in E$, an element $\tilde{A} \in \tilde{E}$ over A is called an **oriented real hypersphere**. The Möbius transformation group contains as its subgroups ones that are isomorphic to the group of congruent transformations of a Euclidean space and ones that are isomorphic to the group of congruent transformations of a non-Euclidean space. That is, the subgroup of $M(n)$ that leaves a point hypersphere invariant is isomorphic to the group generated by congruent transformations and homotheties of the Euclidean space E^n . The subgroup of index 2 (the factor group by a cyclic subgroup of order 2) of the subgroup of $M(n)$ that leaves invariant a real (imaginary) hypersphere is isomorphic to the group of congruent transformation of n -dimensional hyperbolic (elliptic) non-Euclidean space.

In the n -dimensional Euclidean space E^n , consider a hypersphere of radius r with center O . For each point P of E^n , mark a point Q on the ray OP such that $\overline{OP} \cdot \overline{OQ} = r^2$. We call the point transformation that sends P to Q an **inversion** with respect to the hypersphere. A symmetry with respect to a hyperplane, considered as an extreme case of inversions, is also called an inversion. We adjoin a point at infinity to the space E^n to construct an n -dimensional sphere S^n . Each inversion can be extended to a transformation of S^n , which we also call an inversion. Then each Möbius transformation is generated by a finite number of inversions. By a Möbius transformation of S^n , each hypersphere is transformed to a hypersphere. Any angle between two curves that intersect at a point of E^n is invariant under Möbius transformations. Conversely, if $n \geq 3$, each local transformation of E^n that leaves invariant the angle of each pair of intersecting curves is a restriction of a Möbius transformation. However, for $n = 2$ this is not true in general; any transformation that leaves angles invariant is called a conformal mapping. Any Möbius transformation $z \rightarrow w$ on the complex sphere $S^2 = \mathbb{C} \cup \{\infty\}$ can be expressed by an equation of the form $w = (\alpha z + \beta) / (\gamma z + \delta)$ or $w = (\alpha \bar{z} + \beta) / (\gamma \bar{z} + \delta)$, where α, β, γ , and δ are complex numbers such that $\alpha\delta - \beta\gamma \neq 0$ and \bar{z} denotes the complex conjugate of z .

B. Laguerre Geometry

Let Γ be an oriented smooth curve in a Euclidean plane E^2 . The tangent line at a point p of Γ is supplied with an orientation that is induced by the orientation of the curve Γ in an obvious manner. The oriented line l thus obtained is called the **oriented tangent line** of Γ at p .

Let S be the set of oriented lines in E^2 and T be a given set of oriented smooth curves in E^2 . Consider a bijection γ of the direct product $S \times T$ to itself satisfying the following condition: If l is an oriented tangent line of a curve Γ belonging to T and γ sends (l, Γ) to (l', Γ') , then l' is an oriented tangent of the curve Γ' . The set of such bijections forms a group G . Suppose that we have $l \in S$ for which there exist two curves Γ_1 and Γ_2 in T such that l is a common tangent line of Γ_1 and Γ_2 at p_1 and p_2 , respectively. Further suppose that the element $\gamma \in G$ sends (l, Γ_i) to (l', Γ'_i) ($i = 1, 2$). Then the element γ of G is called an **equilong transformation** if the following conditions (i) and (ii) are satisfied: (i) l' is tangent to Γ'_i at points p'_i ($i = 1, 2$), (ii) the distance between p_1 and p_2 is equal to the distance between p'_1 and p'_2 . The set of equilong transformations forms a subgroup H of G . In particular, if T is the set of oriented circles (including point circles), the elements of H are called **Laguerre transformations**. In an obvious manner, we can divide the set of oriented circles into two classes, those with "positive" and those with "negative" orientations. With an oriented circle Γ we associate the pair $\varphi(\Gamma) = (P, r)$, where P is the origin of the circle and r is a real number whose absolute value is equal to the radius of the circle and whose signature coincides with that of the orientation of Γ . The mapping $\varphi: T \rightarrow E^2 \times E^1$ thus defined is called **Lie's minimal projection**. An example of a Laguerre transformation, called a **dilatation**, is given by a bijection γ of $S \times T$ to itself satisfying the following condition: Let $\gamma(l, \Gamma) = (l', \Gamma')$; then l' is parallel to l , the distance between l and l' is a given number, and $\varphi(\Gamma') = (P, r + c)$, where $\varphi(\Gamma) = (P, r)$ and c is a given constant. We note here that the action of a Laguerre transformation $\gamma: (l, \Gamma) \rightarrow (l', \Gamma')$ is determined by its action on S (γ acts on S by $\gamma(l) = l'$). Another example of a Laguerre transformation γ , called a **Laguerre inversion**, is determined by means of a given oriented circle O and a line p that is not tangent to O ; given an oriented line l , its image l' under the action of γ is determined as follows (here we describe the case where l is not parallel to p). There exists a uniquely determined oriented tangent line g of the circle O parallel to l . Now we have a uniquely determined oriented tangent line g' of O that passes through the point of intersection of the lines g and p such that $g' \neq g$. The image l' is the line parallel to g' and passing through the point of intersection of the lines l and p (Fig. 1). Each Laguerre transformation can be written as a product of a finite number of Laguerre inversions. We denote the group of Laguerre transformations by L . The pair (L, S) is, by defini-

tion, a model of **Laguerre geometry**. Notions such as Laguerre inversions, dilatations, and transformations can be generalized to cases of higher dimension by utilizing oriented hyperspheres and oriented hyperplanes.

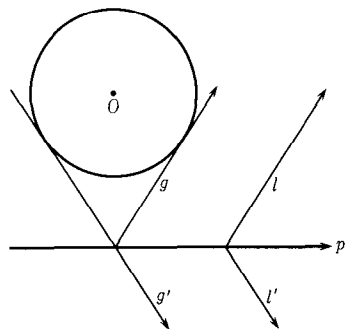


Fig. 1

C. Sphere Geometry

Let S be the set of oriented circles (including point circles and oriented lines) in the Euclidean plane E^2 . Two oriented circles C_1 and C_2 are said to be in contact with each other if they have a point p and an oriented tangent line passing through p in common. (An oriented circle C and an oriented line l are in contact with each other if and only if l is an oriented tangent line of C .) In this case, we call the pair (C_1, C_2) a **contact pair**. A bijection γ of S to itself is called a **Lie transformation** if it sends any contact pair to another. (The Lie transformation is a special case of the **contact transformations**.) An inversion with respect to a circle determines in an obvious manner a Lie transformation, which is also called an **inversion**. We denote the group of Lie transformations by G . Any element γ of G can be written as the product of a finite number of inversions and Laguerre inversions, and G contains the group of Möbius transformations and the group of Laguerre transformations as subgroups. The pair (G, S) is called a model of **circle geometry**. The notion of circle geometry can be generalized to that of **hypersphere geometry** for the case of higher dimensions. Specifically, when we replace E^2 by E^3 and circles by spheres, we have **sphere geometry**.

Let V be a complex 3-dimensional space, and let M, N be the sets of oriented lines and oriented spheres in V , respectively. Then M, N have the structure of 4-dimensional complex manifolds that are homeomorphic to each other. The homeomorphism is given by the **Lie line-sphere transformation** that induces a bijection from the set of pairs of intersecting oriented lines onto the set of pairs of oriented spheres that are in contact with each other.

D. Group-Theoretic Considerations

Here we discuss the preceding three kinds of geometries from the group-theoretic point of view (\rightarrow 137 Erlangen Program). Let us consider the quadratic form Q defined by $Q(x) = -x_0^2 + x_1^2 + x_2^2 + x_3^2 - x_4^2$ in a real projective space P^4 , where $x = (x_0, x_1, x_2, x_3, x_4)$ are homogeneous coordinates. We denote by G the set of all projective transformations of P^4 that leave Q invariant. The group G consists of the set of matrices A of order 5 such that $\det A = 1$ and $Q(Ax) = Q(x)$ holds for all x in P^4 ; we denote by L^3 the set of all points x in P^4 that satisfy $Q(x) = 0$. Then G acts transitively on L^3 . Hence if we denote by H_a the set of all elements of G that fix a point a in L^3 , for example $a = (-1, 1, 0, 0, 0)$, we may assume that $L^3 = G/H_a$. The circle geometry that belongs to the group of Lie transformations of circles is exactly the geometry of the homogeneous space G/H_a . The group G_b of all transformations of G that leave the hyperplane $x_4 = 0$ invariant acts transitively on $L^3 \cap \{x_4 = 0\}$ (G_b is isomorphic to $M(2)$). The geometry of the homogeneous space $G_b/H_a \cap G_b$ is plane conformal geometry. Next, the group G_a of all transformations of G that leave invariant $x_0 + x_1 = 0$ coincides with H_a and acts transitively on $L^3 \cap \{x_0 + x_1 = 0\}$. The geometry of the homogeneous space $L^3 \cap \{x_0 + x_1 = 0\}$ (on which G_a acts transitively) is plane Laguerre geometry. In this sense, the circle geometry that belongs to the group of Lie transformations contains the other circle geometries as subgeometries.

We now describe how plane Laguerre geometry can be realized as the geometry of the space $L^3 \cap \{x_0 + x_1 = 0\}$. Let E^2 be a plane in a Euclidean space E^3 . We fix a Cartesian coordinate system (y_0, y_1, y_2) in E^3 so that E^2 is given by $y_0 = 0$. To each point y of E^3 we can associate an oriented circle in E^2 with center $(0, y_1, y_2)$, radius $|y_0|$, and positive (negative) orientation if y_0 is positive (negative). If y lies on E^2 , the corresponding circle is the point circle y itself. Now let us consider the group G'_a of all affine transformations of E^3 whose rotation parts leave the quadratic form $Q'(y) = -y_0^2 + y_1^2 + y_2^2$ invariant. (G'_a is an isometry with respect to the metric defined by Q' .) Each element of G'_a induces a transformation of the set of oriented circles in E^2 (including the oriented lines and point circles) onto itself. The mappings of E^3 into L^3 defined by $x_0 = (1 + Q'(y))/2$, $x_1 = (1 - Q'(y))/2$, $x_2 = y_1$, $x_3 = y_2$, $x_4 = y_0$ is a one-to-one correspondence of E^3 onto the subset of L^3 such that $x_0 + x_1 \neq 0$. This correspondence induces an isomorphism of G'_a onto G_a . In Laguerre geometry, there are

no essential distinctions between points and oriented circles on E^2 ; and the group G_a acts on a 3-dimensional space of oriented circles (including point circles).

Reference

[1] W. Blaschke, *Vorlesungen über Differentialgeometrie III*, Springer, 1929.

77 (XI.14) Conformal Mappings

A. General Remarks

Let a function $w = f(z)$ that maps a \dagger domain D on the \dagger complex z -sphere homeomorphically onto a domain Δ on the complex w -sphere satisfy the following two conditions: (1) Every curve $C_z: z(t) (0 \leq t \leq 1)$ that starts at any point z_0 in D and possesses a tangent there has an image curve $C_w: w = w(t) = f(z(t)) (0 \leq t \leq 1)$ that also possesses a tangent at the image point $w_0 = f(z_0)$. (2) The angle between any two curves $C_z^{(1)}$ and $C_z^{(2)}$ possessing tangents at z_0 is equal to the angle between their image curves $C_w^{(1)}$ and $C_w^{(2)}$, where the direction of the angle is also taken into account. Then the mapping from D onto Δ is said to be **conformal**, and Δ is called **conformally equivalent** to D . It has been proved that $w = f(z)$ is necessarily a function \dagger analytic in D (D. Men'shov, 1931).

Consequently, the theory of conformal mapping is a branch of the theory of analytic functions. That a function $w = f(z)$ maps a domain D conformally onto a domain Δ means that it is a \dagger meromorphic function \dagger univalent in D and its range is Δ . Then $f'(z) \neq 0$ holds at every (finite) point in D , and the ratio of the lengths of the segments between two points $w_0 = w(0)$, $w(t)$ on C_w and between two points $z_0 = z(0)$, $z(t)$ on C_z tends to a fixed nonvanishing limit $|f'(z_0)|$ as $t \rightarrow 0$ independently of the choice of C_z . Hence if z_1 and z_2 lie on $C_z^{(1)}$ and $C_z^{(2)}$, respectively, and w_1 and w_2 are their image points lying on the image curves $C_w^{(1)}$ and $C_w^{(2)}$, respectively, then the two triangles $\Delta_{z_1 z_0 z_2}$ and $\Delta_{w_1 w_0 w_2}$ are nearly similar in the positive sense, provided that z_1 and z_2 are near enough to z_0 . This justifies the word "conformal," which means "of the same form" (\rightarrow Appendix A, Table 13).

B. Conformal Mapping onto the Unit Disk

A fundamental theorem in the theory of conformal mapping is **Riemann's mapping theo-**

rem, which states that any \dagger simply connected domain D with at least two boundary points can be mapped conformally onto the interior Δ of the unit circle. This theorem is equivalent to the assertion of the existence of \dagger Green's function of D and can be proved in various ways. B. Riemann (1851) gave a proof, based on an idea of C. F. Gauss, by assuming the existence of a solution for a variational problem minimizing the \dagger Dirichlet integral. The logical incompleteness implied by this assumption was later removed by D. Hilbert and others. The proof that is now regarded as simplest is due to L. Fejér and F. Riesz's method (T. Radó, 1922, 1923), which applies \dagger normal family theory. On the other hand, the **osculating process** due to P. Koebe (1912) is a purely constructive method of proving existence that is also applicable to the case of \dagger multiply connected domains. The mapping function $w = f(z)$ in Riemann's mapping theorem is uniquely determined under the normalization condition $f(z_0) = 0$, $\arg f'(z_0) = \theta_0$ at a point z_0 in D , where θ_0 is a given angle.

The other types of simply connected domains are the Riemann sphere less one point and the sphere itself. In both cases every conformal mapping of these domains is a linear transformation. A simply connected domain is called **hyperbolic**, **parabolic**, or **elliptic** if it is conformally equivalent to the unit disk, the complex plane, or the Riemann sphere, respectively.

Let $w = f(z)$ map a simply connected domain D on the z -sphere conformally onto a simply connected domain Δ on the w -sphere. If a sequence $\{z_v\}$ in D tends to a boundary point ζ of D , then the corresponding sequence $\{w_v\}$ ($w_v = f(z_v)$) has no \dagger accumulation point in Δ and does not necessarily tend to a boundary point of Δ .

If for any sequence $\{z_v\}$ tending to ζ , $\{f(z_v)\}$ tends to a unique point ω on the boundary of Δ , it is said that $f(z)$ possesses a **boundary value** ω at ζ . To investigate the behavior of $\{w_v\}$ on a hyperbolic domain D , we can assume, by using a suitable mapping if necessary, that D is bounded. Then the problem is reduced, in view of Riemann's mapping theorem, to the case where D is a bounded simply connected domain and Δ is the unit disk $|w| < 1$.

C. Correspondence between Boundaries

Concerning the correspondence between boundaries under the conformal mapping $w = f(z)$ of a bounded simply connected domain D onto the unit disk $|w| < 1$, we have the following three theorems:

(1) To any \dagger accessible boundary point z_C of D there corresponds a unique point on $|w|=1$, and to any distinct accessible boundary points z_{C_1} and z_{C_2} of D there correspond distinct points on the unit circumference. Furthermore, the set of all points on $|w|=1$ that correspond to accessible boundary points of D has \dagger angular measure equal to 2π .

(2) There is a one-to-one correspondence between \dagger boundary elements of D and points on $|w|=1$ (C. Carathéodory).

(3) Let $w=f(z)$ map the interior D of a \dagger Jordan curve C conformally onto the unit disk $\Delta: |w|<1$. Then it possesses a boundary value, say $f(\zeta)$, at every point ζ on C that satisfies $|f(\zeta)|=1$. Hence $f(z)$ is continuous on the closed domain $\bar{D}=D\cup C$ and maps \bar{D} bijectively onto the closed disk $\bar{\Delta}: |w|\leq 1$. Similarly, the inverse function $z=\varphi(w)$ has an analogous property and maps $\bar{\Delta}$ bijectively and continuously onto \bar{D} ; that is, a conformal mapping of the interior D of a Jordan curve onto the unit disk Δ can be extended into a homeomorphism of the closure \bar{D} to $\bar{\Delta}$ (Carathéodory).

In this case, if the Jordan curve C contains a \dagger regular analytic arc Γ , then the mapping function $w=f(z)$ can be prolonged analytically beyond Γ (except at the endpoints of Γ). Hence the mapping $w=f(z)$ is conformal at interior points of Γ .

Problems on the correspondence of angles at the boundary are closely related to \dagger angular derivatives. These problems have been attacked by Carathéodory, S. Warschawski, J. Wolff, and others.

D. Schwarz-Christoffel Transformation

The problem of determining the form of an analytic function that maps the interior of a circle or a half-plane conformally onto the interior of a polygon was first dealt with by H. A. Schwarz and E. B. Christoffel (1869) (\rightarrow Appendix A, Table 13).

Let P be a polygon in the complex w -plane with vertices b_μ ($\mu=1, \dots, m$) and interior angles $\alpha_\mu\pi$ at b_μ . Then a function $w=f(z)$ that maps a circular disk or a half-plane in the z -plane conformally onto the interior of P is given by

$$f(z) = C \int \prod_{\mu=1}^m (a_\mu - z)^{\alpha_\mu - 1} dz + C'.$$

Here $b_\mu = f(a_\mu)$, and C, C' are constants depending on the position and magnitude of the polygon P . If instead of a disk we consider a half-plane in the z -plane, and if one of the a_μ is the point at infinity, then, modifying this formula by deleting the factor of the integrand corresponding to $a_\mu = \infty$, we obtain a formula

for $w=f(z)$. This representation is called the **Schwarz-Christoffel transformation formula**. This formula was originally derived by Christoffel to solve a problem of 2-dimensional distribution of stationary temperature. It then found extensive application to several problems of conformal mapping concerning polygonal domains and to problems of determining force lines or stream lines and equipotential lines in 2-dimensional electrostatics or hydromechanics.

An analogous formula is also derived for a function that maps the interior of a circle or a half-plane conformally onto the exterior of a polygon. In connection with these formulas, a third-order differential equation that is satisfied by a function mapping the interior of a circle onto a domain bounded by a circular polygon is found useful in the theory of automorphic functions. Moreover, a representation analogous to the Schwarz-Christoffel transformation formula is obtained for a function that maps the interior of a circle onto a curvilinear polygonal domain bounded by arcs of logarithmic spirals with the origin as asymptotic point.

E. Conformal Mapping of Multiply Connected Domains

It is also important to consider problems concerning one-to-one conformal mapping of a multiply connected domain on the z -sphere onto a suitable multiply connected domain \mathfrak{D} on the w -sphere. The two domains D and \mathfrak{D} are then homeomorphic, but the converse is not true; i.e., there does not necessarily exist a one-to-one conformal mapping between D and \mathfrak{D} even when they are homeomorphic. Now let D and \mathfrak{D} be multiply connected domains on the z - and w -planes, respectively, both possessing at least three boundary points. The \dagger universal covering surfaces \hat{D} and $\hat{\mathfrak{D}}$ of D and \mathfrak{D} are hyperbolic and they can be mapped onto the unit disks. The groups of their \dagger covering transformations form the \dagger Fuchsian groups G and \mathfrak{G} . Then in order for D to be mapped one-to-one and conformally onto \mathfrak{D} , it is necessary and sufficient that the group \mathfrak{G} be transformed into the group G by a suitable linear transformation.

To a domain of finite \dagger connectivity having only continua for its boundary components, we can associate **conformal invariants** (namely, **moduli**) expressed by one real parameter in the doubly connected case and by $3n-6$ real parameters in the $n(>2)$ -connected case. A one-to-one conformal mapping is possible only within a class of domains having the same invariants (\rightarrow 416 Teichmüller Spaces).

While a circular disk is taken as a canonical domain in the simply connected case, an annulus is often taken as a canonical domain in the doubly connected case. In the latter case, the logarithm of the ratio (> 1) of the radii of two concentric boundary circles is usually called the **modulus**. There are various types of n (≥ 2)-connected canonical domains, for instance, the whole plane, a circular disk or annulus slit along concentric circular arcs or radial segments, a parallel slit plane, etc. The possibility of a one-to-one conformal mapping of a given domain onto a canonical domain of such a type was proved by Hilbert, Koebe, and others in a potential-theoretic way and by E. Rengel, R. de Possel, H. Grunsky, and others in a purely function-theoretic way. Such canonical domains are characterized by some extremal properties. For example, the horizontal parallel slit mapping function $p_0(z, z_0) = (z - z_0)^{-1} + a[p_0](z - z_0) + \dots$, $z_0 \in D$, has the extremal property that it is the unique function maximizing $\operatorname{Re} a[f]$ within the family of univalent functions $f(z) = (z - z_0)^{-1} + a[f](z - z_0) + \dots$ in D . For a general domain D , $p_0(z, z_0)$, called the extremal horizontal slit mapping, is defined as the limit function of the sequence of horizontal slit mappings $p_0^{(n)}(z, z_0)$ of D_n , $z_0 \in D_n$. Here $\{D_n\}_{n=1}^\infty$ is a canonical exhaustion of D . It has the same extremal property as in the case of finite connectivity. The boundary components of the image domain of D under $p_0(z, z_0)$ consist of a horizontal slit and a point. The parallel slit mapping $p_\theta(z, z_0)$ in the direction of θ is defined similarly to the one that maximizes $\operatorname{Re} e^{-2i\theta} a[f]$ [8].

For a domain of infinite connectivity, by accumulation of boundary components, a pointlike boundary component can be mapped onto a continuum. A boundary component γ of a domain D is called **weak** if its image $f(\gamma)$ is a point under every conformal mapping f on D . γ is called **strong** if $f(\gamma)$ always consists of more than one point. A boundary component γ that is neither weak nor strong is called **unstable** (L. Sario, *J. Analyse Math.*, 5 (1956)). Criteria of weakness, etc., by means of extremal length are given in [8]. Moreover, in terms of extremal length, a generalization of a boundary element is given by Oikawa and Suita [9]. On the other hand, the existence of one-to-one and conformal mapping of a domain bounded by a finite number of curves onto the whole plane with mutually disjoint circular disks removed was proved by Koebe and later derived by J. Douglas and R. Courant as a particular case of the existence of a solution of \dagger Plateau's problem [10]. M. Schiffer showed that the mapping function is a solution of an extremal problem involving

Fredholm eigenvalues [11]. L. Bieberbach and H. Grunsky showed the possibility of mapping an n -connected domain onto an n -sheeted disk. Concerning doubly connected domains, detailed investigations were made by O. Teichmüller, Y. Komatu, and others. For a domain D bounded by n analytic Jordan curves, as an extension of the Schwarz lemma, Ahlfors showed that there exists a unique extremal function $f_0(z)$ maximizing $\operatorname{Re} f'(z_0)$, $z_0 \in D$, within the family of analytic functions $f(z)$ satisfying $|f(z)| < 1$ in D and that $f_0(z)$ maps D onto the n -sheeted unit disk (1947). $f_0(z)$ is called the **Ahlfors function** at z_0 . For a general domain the uniqueness of the extremal function in nontrivial cases was established first by S. Havinson (1961) and later by L. Carleson (1967) and S. Fisher (1969).

Most \dagger kernel functions of a plane domain D have connections with conformal mappings. For example, the Bergman kernel $U(z, \zeta)$ of exact differentials is equal to $(2\pi)^{-1}(p'_0(z, \zeta) - p'_{\pi/2}(z, \zeta))$. The adjoint kernel $V(z, \zeta)$ is defined by $(2\pi)^{-1}(p'_0(z, \zeta) + p'_{\pi/2}(z, \zeta))$. M. Schiffer showed that the integral of $V(z, \zeta)$, $S(z, \zeta) = \int_\zeta^z V(z, \zeta) dz$, is univalent and maps D onto a domain bounded by n analytic convex curves if D is bounded by n Jordan curves [13]. For a general domain D , this is true in the sense that each boundary component of the image domain under $S(z, \zeta)$ is a convex set (K. Oikawa and N. Suita, *Kôdai Math. Sem. Rep.*, 16 (1964)). It is known that $S(z, \zeta)$ maximizes the area of the complementary set of the image domain under $f(z)$ within the family of univalent functions $f(z) = (\pi(z - \zeta))^{-1} + b_1(z - \zeta) + \dots$. The maximum value multiplied by 4π is called the **span** of the domain D ; it is equal to $a[p_0] - a[p_{\pi/2}]$. For the \dagger Szegő kernel function $k(z, \zeta)$ of a domain D bounded by n analytic curves, the adjoint kernel $l(z, \zeta)$ is defined, and the Ahlfors function $f_0(z)$ at ζ is expressed by $k(z, \zeta)/l(z, \zeta)$ (P. Garabedian, *Trans. Amer. Math. Soc.*, 67 (1949)).

F. Universal Constants

Among various universal constants appearing in the theory of conformal mapping, Bloch's constant is especially famous. A. Bloch (1924) showed that a covering surface over the w -plane obtained from a mapping $w = F(z) = z + \dots$ that is one-to-one, conformal, and holomorphic in $|z| < 1$ always contains a \dagger univalent (schlicht) disk whose radius B is a positive number independent of the function F (**Bloch's theorem**). The supremum \mathfrak{B} of such constants B is called **Bloch's constant**. The true value of \mathfrak{B} is yet unknown, but estimates have

been given by Ahlfors (*Trans. Amer. Math. Soc.*, 43 (1938)) and by Ahlfors and Grunsky (*Math. Z.*, 42 (1937)) in the form

$$\frac{\sqrt{3}}{4} \leq \mathfrak{B} \leq \sqrt{\pi} 2^{1/4} \frac{\Gamma(1/3)}{\Gamma(1/4)} \left(\frac{\Gamma(11/12)}{\Gamma(1/12)} \right)^{1/2} = 0.4719 \dots$$

M. Heins showed that the equality sign can be deleted from the left-hand side (*Nagoya Math. J.*, 21 (1962)) and C. Pommerenke gave a simpler proof for the same fact (*J. London Math. Soc.*, 2 (1970)). It is conjectured that the correct value of \mathfrak{B} is equal to the upper bound of Ahlfors and Grunsky. **Landau's constant** \mathfrak{L} corresponding to the case where the image disks are not necessarily univalent satisfies $0.5 \leq \mathfrak{L} < 0.55$. The lower and upper bounds are due to Ahlfors (cited above) and Landau (*Math. Z.*, 30 (1929)). Pommerenke showed $0.5 < \mathfrak{L}$ (1970, cited above). For the family of univalent functions these constants coincide, and the value is called the **schlicht Bloch constant**, denoted by \mathfrak{A} . Its lower and upper bounds are known to be 0.5705 (J. Jenkins, *J. Math. Mech.*, 10 (1961)) and 0.6565 (R. Robinson, *Bull. Amer. Math. Soc.*, 41 (1935)). (For distortion theorems and coefficient problems \rightarrow 438 Univalent and Multivalent Functions.)

References

- [1] L. V. Ahlfors, *Complex analysis*, McGraw-Hill, second edition, 1979.
- [2] Z. Nehari, *Conformal mappings*, McGraw-Hill, 1952.
- [3] D. Gaier, *Konstruktive Methoden der konformen Abbildung*, Springer, 1964.
- [4] C. Carathéodory, *Conformal representation*, Cambridge Univ. Press, 1932, second edition, 1952.
- [5] N. Ohtsuka, *Dirichlet problem, extremal length and prime ends*, Van Nostrand, 1970.
- [6] H. Grunsky, *Lectures on theory of functions in multiply connected domains*, Vandenhoeck & Ruprecht, 1978.
- [7] G. M. Golusin, *Geometric theory of functions of a complex variable*, Amer. Math. Soc. Transl., 1969. (Original in Russian, 1952.)
- [8] L. Sario and K. Oikawa, *Capacity functions*, Springer, 1969.
- [9] K. Oikawa and N. Suita, *On conformal mappings onto incised radial slit discs*, *Kôdai Math. Sem. Rep.*, 22 (1970), 45–72.
- [10] R. Courant, *Dirichlet's principle, conformal mapping, and minimal surfaces*, with an appendix by M. Schiffer, Interscience, 1950.
- [11] M. Schiffer, *Fredholm eigenvalues of multiply connected domains*, *Pacific J. Math.*, 9 (1959), 211–269.

[12] S. Bergman, *The kernel function and conformal mapping*, Amer. Math. Soc. Math. Surveys, 5 (1970).

[13] M. Schiffer, *The span of multiply connected domains*, *Duke Math. J.*, 10 (1943), 209–216.

[14] L. V. Ahlfors, *Conformal invariants*, McGraw-Hill, 1973.

78 (VI.9) Conic Sections

A. General Remarks

Suppose that we are given two straight lines l and m intersecting at V (but not orthogonally) in the 3-dimensional Euclidean space E^3 . By rotating the line m around l , we obtain a surface \mathfrak{F} . We call this surface \mathfrak{F} a **circular cone** with **vertex** V and **axis** l ; a straight line on the surface passing through V is called a **generating line** of \mathfrak{F} .

A section C of \mathfrak{F} by a plane π not passing through V ($C = \pi \cap \mathfrak{F}$) is called a **conic section** (or simply a **conic**). This C is a plane curve on the plane π . The point set $\mathfrak{F} - V$ consists of two †connected components \mathfrak{F}_1 and \mathfrak{F}_2 . Let π_i ($i = 1, 2, 3$) be planes not passing through V . If the conic section $C_1 = \pi_1 \cap \mathfrak{F}$ is †bounded, then C_1 is contained either in \mathfrak{F}_1 or in \mathfrak{F}_2 and is †connected. We call such a C_1 an **ellipse**. When $C_2 = \pi_2 \cap \mathfrak{F}$ is not bounded but is connected, then π_2 is parallel to one of the generating lines of \mathfrak{F} , and C_2 is contained either in \mathfrak{F}_1 or in \mathfrak{F}_2 . We call such a C_2 a **parabola**. When π_3 intersects both of \mathfrak{F}_1 and \mathfrak{F}_2 , then $C_3 = \pi_3 \cap \mathfrak{F}$ has two connected components and is not bounded. We call such a C_3 a **hyperbola**. These three types exhaust all possible types of conic sections. In particular, if the plane π is perpendicular to the axis l , then $C = \pi \cap \mathfrak{F}$ becomes a circle. Thus a circle is a special kind of ellipse.

B. Foci and Directrices

Let $C = \pi \cap \mathfrak{F}_1$ be an ellipse. The Euclidean space E^3 is divided by π into two †half-spaces E_1^3, E_2^3 (two “sides” of π). If we put $\mathfrak{F}_1 \cap E_1^3 = \mathfrak{F}_{11}$, $\mathfrak{F}_1 \cap E_2^3 = \mathfrak{F}_{12}$, we can construct a sphere S that is contained in E_1^3 , tangent to \mathfrak{F}_{11} along a circle K , and tangent to π at a point F . Similarly, we can construct a sphere S' that is in E_2^3 , tangent to \mathfrak{F}_{12} along a circle K' , and tangent to π at a point F' . We call F, F' the **foci** of the ellipse (Fig. 1).

Let κ, κ' be the planes containing K, K' . Straight lines $d = \kappa \cap \pi, d' = \kappa' \cap \pi$ are called **directrices** of C . Unless C is a circle, we have $F \neq F'$, and κ, π (and κ', π) actually intersect; hence d, d' exist. When $C = \pi \cap \mathfrak{F}$ is a parabola or a hyperbola (Figs. 2 and 3), we can similarly define foci (a parabola has only one focus, F , while a hyperbola has two foci, F, F') and directrices (a parabola has only one directrix, d , while a hyperbola has two directrices, d, d').

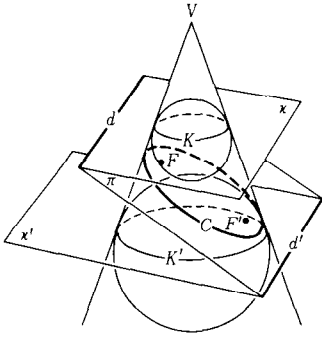


Fig. 1

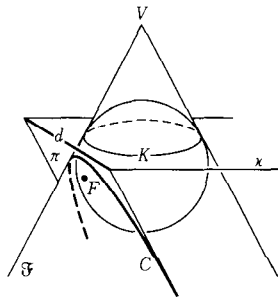


Fig. 2

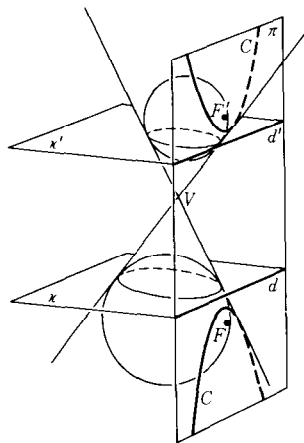


Fig. 3

Let X be a point on the plane π , let $D_F(X)$ be the distance between the point X and a focus F , and let $D_d(X)$ be the distance between X and a directrix d . Then the curve C is the locus of the points X satisfying the condition $D_F(X) = e \cdot D_d(X)$, where e is a constant. We call

e the **eccentricity** of the conic section C . According as C is an ellipse, a parabola, or a hyperbola, we have $e < 1$, $e = 1$, or $e > 1$. A circle is an ellipse whose eccentricity is zero. An ellipse is also characterized as the locus of X such that $FX + F'X = 2a$; a hyperbola is the locus of X such that $|FX - F'X| = 2a$, where a is a positive constant. When there are two foci, the straight line FF' is perpendicular to directrices d and d' .

C. Canonical Forms of Equations

When C is a hyperbola or an ellipse that is not a circle, C has two foci, F and F' . In this case, the midpoint O of the segment FF' is the center of symmetry of C (when C is a circle, its center O is, of course, the center of symmetry of C). We call O the **center** of C ; an ellipse or a hyperbola is called a **central conic**. If we choose a rectangular coordinate system (x, y) having O as the origin and FF' as x -axis, then the equation of C can be expressed in the form

$$x^2/a^2 \pm y^2/b^2 = 1, \quad a, b > 0. \quad (1)$$

According as C is an ellipse or a hyperbola, we take the $+$ or $-$ of the double sign. If C is an ellipse, we have $a > b$. Furthermore, $e = \sqrt{a^2 - b^2}/a$ if C is an ellipse and $e = \sqrt{a^2 + b^2}/a$ if C is a hyperbola. We also have $F = (ae, 0)$ and $F' = (-ae, 0)$; the equations of directrices are $x = \pm a/e$.

On the other hand, if C is a parabola, the straight line that is perpendicular to the directrix d and passes through F becomes the axis of symmetry of C . We call this straight line the **axis** of C ; the intersection O of the axis and C is called the **vertex** of C . If we choose a rectangular coordinate system (x, y) having O as the origin and having the axis of C as the x -axis, the equation of C can be expressed in the form

$$y^2 = 4ax, \quad a > 0. \quad (2)$$

We call (1) and (2) the **canonical** (or **standard**) forms of the equation of C . We call the associated coordinate system the **canonical coordinate system**. Suppose that C is an ellipse (hence $a > b$). Let A, A' be points of intersection of the x -axis and the ellipse and B, B' be the points of intersection of the y -axis and the ellipse. We call AA' the **major axis** of C and BB' the **minor axis** of C . If C is a hyperbola and (x, y) is the canonical coordinate system, we call the x -axis the **transverse axis** and the y -axis the **conjugate axis**. If C is a central conic, the x - and y -axes of the canonical coordinate system are called the **principal axes**; if C is a parabola, the x -axis is sometimes called the **principal axis**.

D. Properties of Ellipses

An ellipse may be considered the image of a circle under a \dagger parallel projection. Consequently, the section of a circular cylinder by a plane is an ellipse. Also, if we are given a circle C and a fixed diameter D of C , an ellipse is obtained as the locus of the points X lying on lines PM , which are perpendicular to D , with $P \in C$, $M \in D$, satisfying the condition that the ratio $PM:XM$ is constant.

Suppose that we are given two concentric circles having the center at the origin O and with radii a , b . Let P , Q be points of intersection of moving half-lines through O and the two circles. Then the locus of points X of intersection of the ordinates (lines parallel to the y -axis) passing through P and the abscissae (lines parallel to the x -axis) passing through Q is an ellipse (Fig. 4). Suppose that the equation of an ellipse is given by $x^2/a^2 + y^2/b^2 = 1$, with $a > b$. Then the lengths of its major axis and minor axis are $2a$ and $2b$, respectively.

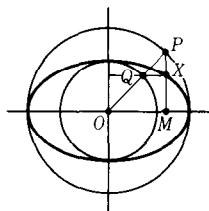


Fig. 4

Given an ellipse and its center O , the circle with center O and diameter equal to the major axis of the ellipse is called the **auxiliary circle** of the ellipse. Given an ellipse C and its focus F , the auxiliary circle of C is the locus of the points X satisfying the condition that the line FX is perpendicular to a tangent line to C passing through X . Suppose that X is a point on an ellipse with foci F , F' . Let TT' be the line tangent to the ellipse at X (X lies between T and T'). Then the angle $\angle TXF'$ is equal to $\angle T'XF$ (Fig. 5). Consequently, the rays starting from one focus of an ellipse and “reflected” by the ellipse converge on the other focus of the ellipse. Also, the product of the distances from two foci of an ellipse to an arbitrary tangent is constant and is equal to b^2 .

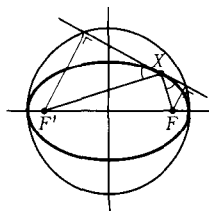


Fig. 5

The ellipse $C: x^2/a^2 + y^2/b^2 = 1$ is expressed parametrically in the form

$$x = a \cos \theta, \quad y = b \sin \theta. \quad (3)$$

We call the parameter θ the **eccentric angle** of a point (x, y) on C . Consequently, C is a \dagger Jordan curve and divides the plane into two parts, the inside and the outside. The inside is the set of points (x, y) satisfying $x^2/a^2 + y^2/b^2 < 1$, and the outside is the set of points (x, y) satisfying $x^2/a^2 + y^2/b^2 > 1$. The inside is a \dagger convex set.

From a point Q outside C , two tangents to C can be drawn. The locus of points Q such that these two tangents are orthogonal is the circle $x^2 + y^2 = a^2 + b^2$. We call this circle the **direct circle**. The area of the “sector” OAX formed by two points $A(a, 0)$, $X(a \cos \theta, b \sin \theta)$ ($\theta > 0$) and the origin O is $ab\theta/2 = (ab/2) \text{Arcos}(x/a)$; the length of the arc AX of the ellipse is represented by the value of the \dagger elliptic integral

$$a \int_0^\theta \sqrt{1 - e^2 \cos^2 \theta} d\theta = aE(\pi/2 - \theta, e).$$

In particular, the area inside an ellipse is equal to πab , and the whole length of the ellipse is $4aE(0, e)$.

With respect to a polar coordinate system (r, θ) having the focus $F(ae, 0)$ as the origin and the ray directed positively along the x -axis as the initial line, the equation of the ellipse C is

$$r = \frac{l}{1 + l \cos \varphi}, \quad l = \frac{b^2}{a}. \quad (4)$$

Here l is equal to half of the length of the chord that is perpendicular to the major axis and passes through the focus. (This chord is called the **latus rectum** of the ellipse.) Suppose that F is a fixed point and that X is a moving particle attracted toward F by a \dagger central force inversely proportional to the square of the length of FX . Suppose further that X begins with an initial velocity whose direction is tangent to the ellipse C with focus F . Then X always moves on C , and the areal velocity described by the radius FX is constant (Kepler's second law).

E. Properties of Hyperbolas

Two straight lines $x^2/a^2 - y^2/b^2 = 0$, that is, $y/x = \pm b/a$, are \dagger asymptotes of the hyperbola $C: x^2/a^2 - y^2/b^2 = 1$. The hyperbola $C': x^2/a^2 - y^2/b^2 = -1$ is called the **conjugate hyperbola** of C . When $a = b$, the asymptotes are orthogonal to each other, and C , C' are congruent. In this case, we call C a **rectangular hyperbola** (or **equilateral hyperbola**). When we draw parallels to asymptotes from a point X on C , the area of the parallelogram formed by these

lines and two asymptotes is constant. (In particular, when C is a rectangular hyperbola, the equation of C becomes $xy = k^2/2$ if we take two asymptotes as coordinate axes. The segment cut off by the asymptotes on the tangent to C at X is divided equally at X . In the case of the hyperbola as well, the product of the distances from two foci to an arbitrary tangent is constant and is equal to b^2 , and the angle between two straight lines joining two foci to a point X on C is divided equally by the tangent at X .

A hyperbola C is represented parametrically by

$$x = a \sec \theta, \quad y = b \tan \theta. \quad (3')$$

In this case also, we call θ the **eccentric angle** of (x, y) . If we use the hyperbolic functions and the parameter u , the equation of a hyperbola can be written as

$$x = a \cosh u, \quad y = b \sinh u \quad (3'')$$

instead of (3'). The area of a "sector" OAX formed by two points $A(a, 0)$, $X(x, y)$ on the hyperbola and the origin O is, in this case,

$$\frac{abu}{2} = \frac{ab}{2} \operatorname{Arccosh} \frac{x}{a} = \frac{ab}{2} \log \frac{x + \sqrt{x^2 - a^2}}{a}.$$

The length of the arc \widehat{AX} of the hyperbola is given by the elliptic integral

$$\int_0^x \sqrt{\frac{e^2 x^2 - a^2}{x^2 - a^2}} dx.$$

With respect to a polar coordinate system (r, φ) having the focus $F(ae, 0)$ as origin and the ray directed positively along the x -axis as the initial line, the equation of the hyperbola C becomes

$$r = \frac{l}{1 - e \cos \varphi}, \quad l = \frac{b^2}{a}, \quad (4')$$

where l is equal to half the length of the chord passing through the focus and perpendicular to the principal axis. (This chord, too, is called the **latus rectum**.)

F. Properties of Parabolas

The curve described by a particle attracted by "gravitation" in a fixed direction and affected by no other force is a parabola (G. Galilei). The tangent at a point X on a parabola makes equal angles with the straight line joining F and X and the direction of the principal axis (Fig. 6). Consequently, if the rays starting from the focus of a parabola are "reflected" by the parabola, they all become rays parallel to the principal axis. Let X' be

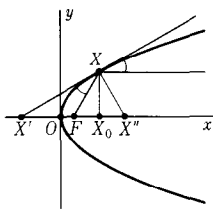


Fig. 6

the point of intersection of the tangent at $X(x_0, y_0)$ and the x -axis, X'' the point of intersection of the normal at X and the x -axis, and X_0 the foot of the perpendicular from X to the x -axis. Then $FX = FX'$, $\triangle FXX'$ is an isosceles triangle, and XX' is divided equally by the y -axis. Consequently, the locus of the foot of a perpendicular from F to a tangent is the y -axis. Also, the length of subtangent $X'X_0 = 2x_0$, and the length of subnormal $X''X_0 = 2a = 2OF$. Conversely, a curve whose length of subnormal is constant is a parabola. The locus of the midpoints of parallel chords of a parabola is a straight line parallel to the principal axis.

With respect to a polar coordinate system having the focus as origin and the ray directed positively along the x -axis as the initial line, the equation of a parabola is

$$r = \frac{l}{1 - \cos \varphi}, \quad l = 2a. \quad (4'')$$

The area bounded by a chord BC and an arc \widehat{BC} of a parabola (Fig. 7) is equal to $4/3$ the area of $\triangle ABC$, where A is the point of contact on the tangent of the parabola parallel to BC (Archimedes). Also, the length of the arc \widehat{OX} of parabola (2) is

$$\frac{y_0}{4a} \sqrt{y_0^2 + a^2} + a \log \frac{y_0 + \sqrt{y_0^2 + a^2}}{2a},$$

X having the coordinates (x_0, y_0) .

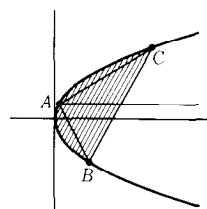


Fig. 7

G. Conjugate Diameters

The **diameter** is a straight line passing through the center of a central conic. The locus of the midpoints of the chords parallel to a diameter d is another diameter d' , called **conjugate** to d .

Then the diameter conjugate to d' is d . The x -axis and y -axis of a canonical coordinate system form a set of conjugate diameters. Let $2a'$, $2b'$ be the lengths of the segments (sometimes called conjugate diameters) cut off from d , d' by the curve or by the curve and the conjugate one for hyperbolas and by ω , the angle between d and d' . Then the following relations hold (as to the double signs \pm , we take $+$ in the case of an ellipse and $-$ in the case of a hyperbola): $a'^2 \pm b'^2 = a^2 \pm b^2$, $a'b' \sin \omega = ab$. The product of the slopes of d , d' is equal to $\pm b^2/a^2$. With respect to an oblique coordinate system having d and d' as axes, the equation of the curve is $x^2/a'^2 \pm y^2/b'^2 = 1$.

H. Confocal Conic Sections

The set of ellipses and hyperbolas having two fixed points F , F' as foci is called the **family of confocal central conics** with foci F and F' (Fig. 8). The family of confocal central conics containing the ellipse $x^2/a^2 + y^2/b^2 = 1$ is represented parametrically by

$$\frac{x^2}{a^2 + \lambda} + \frac{y^2}{b^2 + \lambda} = 1.$$

There exist only one ellipse and only one hyperbola that pass through a point inside each quadrant (for example, a point (x_0, y_0) , $x_0 > 0$, $y_0 > 0$, inside the first quadrant) and belong to the family of curves. The ellipses and hyperbolas of the same family cut each other orthogonally. Thus parameters corresponding to ellipses and hyperbolas belonging to a family of confocal central conics define an orthogonal curvilinear coordinate system, called an \dagger elliptic coordinate system (\rightarrow 90 Coordinates).

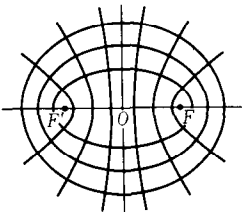


Fig. 8

The set of parabolas having a fixed point F as focus and a straight line passing through F as axis is called a **family of confocal parabolas** (Fig. 9). The family of confocal parabolas containing $y^2 = 4ax$ is the set of curves

$$y^2 = 4(a + \lambda)(x + \lambda).$$

Such families also give rise to orthogonal curvilinear coordinate systems.

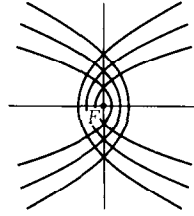


Fig. 9

I. Curves of the Second Order

With respect to a rectangular coordinate system, a curve represented by an equation with real coefficients of the second degree with two variables x , y ,

$$ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0, \quad (5)$$

where $(a, h, b) \neq (0, 0, 0)$, is called a **curve of the second order**. A curve of the second order is either an empty set, one point, one or two straight lines, or a conic section. For equation (5), we put

$$D_0 = \begin{vmatrix} a & h \\ h & b \end{vmatrix}, \quad D = \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} \quad (6)$$

and call D the **discriminant** of the curve of the second order. If $D_0 \neq 0$, $D \neq 0$, and the curve is not an empty set, then the curve is a central conic. If $D_0 > 0$, then the curve is an ellipse or an empty set. If $D_0 < 0$, then the curve is a hyperbola. If $D_0 = 0$, $D \neq 0$, then the curve is a parabola. If $D = 0$, $D_0 > 0$, then the curve consists of one point. If $D = 0$, $D_0 < 0$, then the curve is two intersecting straight lines. If $D = D_0 = 0$, then the curve is an empty set, one straight line, or two parallel straight lines.

J. Poles and Polars

Let $F(x, y) = ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$ be the equation of a conic C and (x_0, y_0) the coordinates of a point P on the plane. A straight line P^* having the equation

$$ax_0x + h(x_0y + xy_0) + by_0y + g(x + x_0) + f(y + y_0) + c = 0$$

is called the **polar** of P with respect to C (Fig. 10). When the polar of a point P is l , we call P the **pole** of l and denote it by l^* . In general, l^* is uniquely determined by l , and $P^{**} = P$, $l^{**} = l$. If $Q \in P^*$, then $P \in Q^*$. If $P' \in l$, then $l^* \in P'^*$. When a straight line passing through P intersects C at X , Y and intersects P^* at P' , then P , P' are \dagger harmonic conjugate with respect to X , Y . In particular, if $P \in P^*$, then $P \in C$, and P^* becomes the tangent of C at P . Given a

triangle $\triangle PQR$ on the plane of C , we call the triangle with sides P^* , Q^* , R^* the **polar triangle** of $\triangle PQR$. Let $Q^* \cap R^* = P'$, $R^* \cap P^* = Q'$, and $P^* \cap Q^* = R'$. Then the three straight lines $P \cup P'$, $Q \cup Q'$, $R \cup R'$ meet at a point (M. Chasles). When the polar triangle of $\triangle PQR$ coincides with itself, then $\triangle PQR$ is called a **self-polar triangle**. The polar of a focus is a directrix.

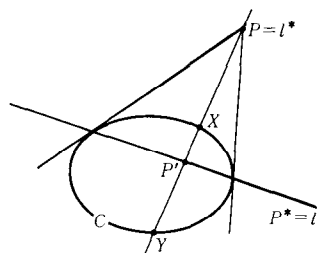


Fig. 10

K. Curves of the Second Class

When the coefficients u, v, w of a straight line $ux + vy + w = 0$ satisfy an equation with real coefficients of the second order,

$$Au^2 + 2Huv + Bv^2 + 2Guw + 2Fvw + Cw^2 = 0, \quad (5')$$

where $(A, H, B) \neq (0, 0, 0)$, the curve enveloped by these straight lines is called a **curve of the second class**. Let Δ be the discriminant defined analogously to D in (6) by using A, H, B, \dots instead of a, h, b, \dots . A curve of the second class (with $\Delta \neq 0$) is essentially the same as a curve of the second order with $D \neq 0$. In order for the curve (5') with $\Delta \neq 0$ to coincide with the curve (5) with $D \neq 0$, it is necessary and sufficient that A, B, C, F, G, H be proportional to the \dagger cofactors of a, b, c, f, g, h in the determinant D given by (6). If $\Delta = 0$, then (5') represents either the empty set, or a point (regarded as the set of straight lines passing through the point), or two points.

From a projective point of view, a curve of the second order is defined as a locus of the point of intersection $l \cap l' = X$ of corresponding lines l and l' when two \dagger pencils of lines $A(l, m, \dots)$, $A'(l', m', \dots)$ passing through two different centers A and A' are in correspondence under a \dagger projective mapping f (J. Steiner) (Fig. 11). From this it can be proved that three points of intersection of three pairs of opposite sides (AB, DE) , (BC, EF) , (CD, FA) of a hexagon inscribed in a curve of the second order are on the same straight line (**Pascal's theorem**, Fig. 12). In particular, if the curve of the second order in this theorem consists of two straight lines, the theorem coincides with **Pappus's theorem** (Fig. 13). We call this

straight line l the **Pascal line** of $ABCDEF$.

Given a set of six points A, B, C, D, E, F on a curve of the second order, by considering all possible combinations of the points, we get 60 Pascal lines. A configuration consisting of these 60 lines is called **Pascal's configuration**, and has been studied by Steiner, Kirkman, and others. As a \dagger dual to Pascal's theorem, **Brianchon's theorem** holds: Three diagonals of a hexagon with a curve of the second class inscribed meet at a point (Fig. 14).

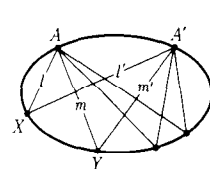


Fig. 11

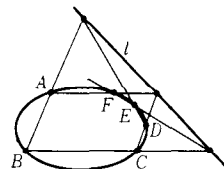


Fig. 12

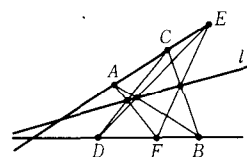


Fig. 13

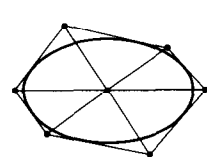


Fig. 14

References

- [1] G. Salmon, A treatise on conic sections, Longmans, Green, sixth edition, 1879 (Chelsea 1962).
- [2] H. F. Baker, Principles of geometry II. Plane geometry, Cambridge Univ. Press, 1922.
- [3] M. Protter and C. Morrey, Jr., Calculus with analytic geometry, Addison-Wesley, 1962.

79 (II.20) Connectedness

A. General Remarks

A \dagger topological space X is said to be **connected** if there are no proper closed subsets A and B of X such that $A \cap B = \emptyset$ and $A \cup B = X$ (C. Jordan, *Cours d'Analyse* I, 1883). A subset S of X is **connected** if S considered as a \dagger subspace of X is connected.

If a subset S of X is connected, then the \dagger closure \bar{S} is also connected. Let $\{A_\alpha\}$ be a family of connected subsets of X such that either $\bar{A}_\alpha \cap A_\beta \neq \emptyset$ or $A_\alpha \cap \bar{A}_\beta \neq \emptyset$ for any pair A_α and A_β . Then the union $\bigcup_\alpha A_\alpha$ is connected. The continuous image of a connected set is connected. The \dagger product space $\prod_\alpha X_\alpha$ of a family of connected spaces $\{X_\alpha\}$ is also con-

nected. Let $\{A_\alpha\}$ be a family of connected subsets of X and A_0 be a connected subset of X . If $A_0 \cap A_\alpha \neq \emptyset$ for every A_α , then the union $A_0 \cup \bigcup_\alpha A_\alpha$ is also connected. For a point p of a topological space X , the union of all connected subsets containing p is connected and is called the **connected component** of p (F. Hausdorff, 1927).

The n -dimensional Euclidean space E^n ($n \geq 0$) and the n -dimensional unit sphere S^n ($n \geq 1$) are connected, whereas S^0 , consisting of two points, is not connected. $E^n \setminus \{0\}$ is connected for $n \geq 2$, whereas $E^1 \setminus \{0\}$ is not connected. This fact implies that E^1 is not homeomorphic to E^n ($n \geq 2$). A connected open subset of a topological space X is called a **domain** (or **region**) in X .

A topological space X is said to be **locally connected** at a point p if for every open set U containing p , there is an open set V containing p and contained in the connected component of p in U . X is said to be **locally connected** if it is locally connected at each point of X . A space X is locally connected if and only if every connected component of every open subset is open in X .

There are connected spaces that are not locally connected. For example, the **comb space** $\{(x, y) \in \mathbb{R}^2 \mid x = 1/n \text{ and } 0 \leq y \leq 1 \text{ (} n = 1, 2, 3, \dots) \text{ or } 0 \leq x \leq 1 \text{ and } y = 0\}$ is not locally connected at $(0, 1)$ (see Fig. 1).

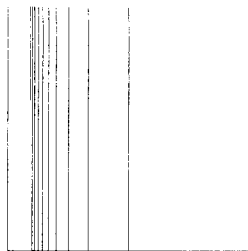


Fig. 1
Comb space.

B. Arcwise Connectedness

Two points a and b of a topological space X are said to be **joined by an arc** in X if there is a continuous map $f(t)$ of the closed interval $I = [0, 1]$ into X such that $f(0) = a$ and $f(1) = b$. A topological space X is said to be **arcwise connected** (or **path-connected**) if every two points of X are joined by an arc in X . For a point p of X , the union of all arcwise-connected subsets in X containing p is arcwise connected and is called the **arcwise-connected component** (or **path-component**) of p . A topological space X is said to be **locally arcwise connected** at a point p if for every open subset U containing p , there is an open subset V

containing p such that every two points in V are joined by an arc in U . X is said to be **locally arcwise connected** if it is locally arcwise connected at each point of X .

An arcwise-connected space is connected, but the converse is not true. The \dagger sinusoid $\{(x, y) \in \mathbb{R}^2 \mid y = \sin 1/x \text{ and } 0 < x \leq 1 \text{ or } x = 0 \text{ and } -1 \leq y \leq 1\}$ is connected but not arcwise connected. A complete metric space is arcwise connected if it is connected and locally connected.

C. Simple Connectedness and n -Connectedness

Denote by S^n the n -dimensional unit sphere and by D^{n+1} the $(n+1)$ -dimensional unit disk. A topological space X is said to be **n -connected** if every continuous map f from S^m to X is extendable over D^{m+1} ($m = 0, 1, 2, \dots, n$). 0-Connectedness is equivalent to arcwise connectedness. A 1-connected space is also called **simply connected**. (For simple connectedness and n -ply connectedness of plane domains → 333 Plane Domains) A topological space X is said to be **locally n -connected** at a point p if for every open subset U containing p , there is an open subset V containing p and contained in U such that any continuous map $f: S^m \rightarrow V$ is extendable to a continuous map $\tilde{f}: D^{m+1} \rightarrow U$ ($m = 0, 1, 2, \dots, n$). X is said to be **locally n -connected** if it is locally n -connected at each point of X . A space X is said to be **ω -connected** (or **locally ω -connected**) if X is n -connected (or locally n -connected) for every n .

S^n is $(n-1)$ -connected but not n -connected. Similarly for $E^{n+1} \setminus \{0\}$. If X is n -connected ($n \geq 1$), then the \dagger suspension SX is $(n+1)$ -connected and the \dagger loop space ΩX is $(n-1)$ -connected. The **Hawaiian earring** $\{(x, y) \in \mathbb{R}^2 \mid (x - 1/n)^2 + y^2 = 1/n^2 \text{ (} n = 1, 2, 3, \dots) \}$ is arcwise connected but not locally simply connected. The \dagger cone over the Hawaiian earring is simply connected, but not locally simply connected (Fig. 2).

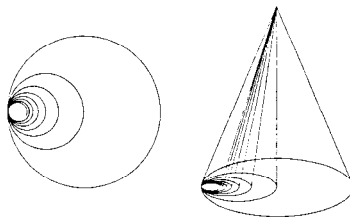


Fig. 2
Hawaiian earring and its cone.

A topological space X is said to be **contractible** if the identity map 1_X is \dagger homotopic to a constant map c_{x_0} to a point x_0 of X (K. Borsuk, *Fund. Math.*, 24, 1935). A topological space X is said to be **locally contractible** at a

point p of X , if for every open subset U containing p , there is an open subset V containing p and contained in U such that the inclusion $i: V \subset U$ is homotopic to a constant map to a point in U . X is said to be **locally contractible** if it is locally contractible at each point of X .

A contractible or locally contractible space is ω -connected or locally ω -connected, respectively. The n -dimensional simplex, the n -dimensional disk D^n , and the n -dimensional Euclidean space E^n are contractible. More generally, a \ast convex set in E^n is contractible. \ast Topological manifolds and \ast polyhedra are locally contractible.

D. Continua and Discontinua

A topological space is said to be **totally disconnected** if each connected component consists of one point. Divide the closed interval $I = [0, 1]$ into three equal parts, and let I_{11} and I_{12} be the closed intervals obtained from I by removing the middle open interval. As the next step, divide I_{11} and I_{12} into three equal parts, respectively, and remove the middle open intervals. Inductively we obtain 2^{n+1} closed intervals $I_{n+1,i}$ ($i = 1, \dots, 2^{n+1}$) from 2^n closed intervals $I_{n,j}$ by removing open intervals lying in their middles. Let $C^{(n)} = \bigcup_{i=1}^{2^n} I_{n,i}$ and $C = \bigcap_{n=1}^{\infty} C^{(n)}$. Then C is called the **Cantor discontinuum** or simply the **Cantor set** or the **ternary set** (G. Cantor, *Math. Ann.*, 21 (1883)) (Fig. 3).



Fig. 3
Cantor discontinuum.

C is a subset of I consisting of points with coordinates $t = (n_1/3) + (n_2/3^2) + \dots + (n_i/3^i) + \dots$, where $n_i = 0$ or 2 . As a topological space C is homeomorphic to the Cartesian product of countably many copies of the discrete space $D = \{0, 1\}$. C has the power of the continuum c and is a compact, totally disconnected, \ast perfect set. The Cartesian product of infinitely many (countably or not) copies of D is called the **general Cantor set**. The continuous image of a general Cantor set is called a **dyadic compactum**. Compact metric spaces and compact groups are examples of dyadic compacta.

A **continuum** is by definition a connected compact metric space consisting of more than one point. A metric space X is said to be **well-chained** if for every two points a, b and $\varepsilon > 0$ there are points x_1, x_2, \dots, x_{n-1} such that $d(x_i, x_{i+1}) < \varepsilon$ ($x_0 = a, x_n = b$). A well-chained compact metric space is a continuum.

Let K be a continuum containing two

points a and b . Then K is said to be **irreducible** between a and b if there is no proper subcontinuum of K containing a and b (L. Zoratti, *Ann. Sci. Ecole Norm. Sup.*, 26 (1909)) (\rightarrow 93 Curves).

A continuum K is said to be **indecomposable** if there are no proper subcontinua K_1, K_2 such that $K = K_1 \cup K_2$ (L. E. J. Brouwer, *Math. Ann.*, 66 (1910)). Simple examples of indecomposable continua have been given by A. Denjoy, *C. R. Acad. Sci. Paris*, 151 (1910); K. Yoneyama, *Tôhoku Math. J.*, 12 (1917); and B. Knaster, *Fund. Math.*, 3 (1922).

C. Kuratowski conjectured the following: If a plane continuum K is homogeneous (that is, for any two points $a, b \in K$ there exists a homeomorphism $h: K \rightarrow K$ with $h(a) = b$) then K is homeomorphic to the circle. A counterexample for this conjecture has been found by R. H. Bing and E. E. Moise. It is an indecomposable, homogeneous, plane continuum and is called the **pseudo-arc** [8].

References

- [1] K. Kuratowski, *Topologie II*, Monograf. Mat. (1950).
 - [2] S. Lefschetz, *Topics in topology*, Ann. Math. Studies, Princeton Univ. Press, 1942.
 - [3] R. L. Wilder, *Topology of manifolds*, Amer. Math. Soc. Colloq. Publ., 1949.
 - [4] G. T. Whyburn, *Analytic topology*, Amer. Math. Soc. Colloq. Publ., 1942.
 - [5] S. T. Hu, *Elements of general topology*, Holden-Day, 1964.
 - [6] O. Hanner, Retraction and extension of mappings of metric and non-metric spaces, *Ark. Mat.*, 2 (1952), 315–360.
 - [7] K. Borsuk, *Theory of retracts*, Monograf. Mat. (1967).
 - [8] J. G. Hocking and G. S. Young, *Topology*, Addison-Wesley, 1961.
- Also \rightarrow references to 425 Topological Spaces.

80 (VII.4) Connections

A. History

The geometric notion of connections originated with T. Levi-Civita's parallelism (*Rend. Circ. Mat. Palermo*, 42 (1917)) and was later generalized to the notion of connections of differentiable fiber bundles. Notions such as affine connections, Riemannian connections, projective connections, and conformal connections can be described in terms of bundles

constructed from the tangent bundles of differentiable manifolds. They are also standard examples of the Cartan connections formulated by E. Cartan and C. Ehresmann.

B. Connections in Principal Bundles

Let $P = (P, \pi, M, G)$ be a differentiable \dagger principal fiber bundle. (For the sake of convenience, we assume that differentiability always means that of class C^∞ .) The total space P and the base space M are \dagger differentiable manifolds, and the projection π is a differentiable mapping. The \dagger structure group G is a \dagger Lie group and acts on P from the right as a transformation group. On each fiber, G acts transitively without fixed points. For elements a, x in G, P , we write $R_a(x) = xa$. The mappings induced on \dagger tangent vector spaces by R_a and π will be denoted by the same letters, namely $R_a: T_x(P) \rightarrow T_{xa}(P), \pi: T_x(P) \rightarrow T_{\pi(x)}(M)$. The tangent vector space $T_x(P)$ at each point x of P is mapped by the projection π onto the tangent vector space $T_p(M)$ at the point $p = \pi(x)$ of M . The kernel of this mapping is denoted by $V_x(P)$, and each vector in $V_x(P)$ is said to be **vertical**. The kernel $V_x(P)$ is the totality of elements of $T_x(P)$ that are tangent to the fiber.

C. Connections

We say that a **connection** is given in P if for each point $x \in P$, a subspace Q_x of the tangent space $T_x(P)$ is given in such a way that the following three conditions are satisfied: (i) $T_x(P) = V_x(P) + Q_x$ (direct sum); (ii) $R_a(Q_x) = Q_{xa}$ (Q is invariant under G); and (iii) the mapping $x \rightarrow Q_x$ is differentiable. A vector in Q_x is said to be **horizontal**.

Now suppose that X is an arbitrary \dagger vector field on P . By condition (i), the value X_x of X at each point x of P can be expressed uniquely as $X_x = Y_x + Z_x$, where $Y_x \in V_x(P)$ and $Z_x \in Q_x$. The vector fields Y and Z defined by Y_x and Z_x ($x \in P$) are called the **vertical** and **horizontal components** of X , respectively. Condition (iii) implies that if X is a differentiable vector field, then its horizontal and vertical components are also differentiable vector fields. Let X be a vector field on the base space M . Since π defines an isomorphism of Q_x and $T_p(M)$ ($p = \pi(x)$), we have a unique vector field X^* on P such that (a) $\pi(X^*) = X$ and (b) $X_x^* \in Q_x$. We call X^* the **lift** of X , and it is invariant under G by condition (ii).

Suppose that a connection is given in P . If C is a piecewise differentiable curve in the base space M , we can define a mapping φ that maps the fiber over the initial point p of C

onto the fiber over the endpoint q of C as follows: Take an arbitrary point x on the fiber at p . Then we have a unique curve C_x^* in P starting at x such that (a) $\pi(C_x^*) = C$, and (b) each tangent vector to C_x^* is horizontal. (C_x^* is called a **lift** of C that starts at x .) The endpoint y of the curve C_x^* belongs to the fiber over q . We set $\varphi(x) = y$. Because $C_{xa}^* = R_a(C_x^*)$, the mapping φ commutes with transformations of G . We call this mapping φ the **parallel displacement** or **parallel translation** along the curve C .

D. Holonomy Groups

Fix a point p in the base space. If C is a closed curve in M starting from p , the parallel displacement along C maps the fiber over p onto itself. So if we fix a point x on the fiber over p , x is transformed by the parallel displacement to a point xa ($a \in G$). Thus each closed curve C starting from p determines an element $a(x, C)$ of G . If C varies over the set of closed curves that start from p , the totality of such elements of G forms a subgroup of G . This subgroup is called the **holonomy group** of the connection defined over P with the reference point x . If M is connected, holonomy groups with different reference points are conjugate. In the above, if we choose as the closed curves C starting from p only those curves that are null-homotopic, the elements $a(x, C)$ form a subgroup of the holonomy group. This is called the **restricted holonomy group**. The holonomy group is a \dagger Lie subgroup of the structure group, and its connected component containing the identity coincides with the restricted holonomy group. Holonomy groups are useful in the study of the behavior of connections.

E. Connection Forms

Let \mathfrak{g} be the Lie algebra (\rightarrow 249 Lie Groups) of the structure group of G of a principal fiber bundle $P = (P, \pi, M, G)$. For each A in \mathfrak{g} , the 1-parameter subgroup $\exp tA$ ($-\infty < t < \infty$) of G defines a \dagger one-parameter group $R_{\exp tA}$ of transformations on P , and it determines a vector field A^* on P (\rightarrow 105 Differentiable Manifolds). Each element of the vector field A^* is vertical at each point x on P , and the A^* ($A \in \mathfrak{g}$) at x generate $V_x(P)$. Moreover, for each element a of G we have $R_a(A^*) = (\text{ad}(a^{-1})A)^*$.

For a connection in P , we define the **connection form** ω on P with values in \mathfrak{g} by the following: (i) $\omega_x(A_x^*) = A$ ($A \in \mathfrak{g}$), and (ii) $\omega_x(X) = 0$ ($X \in Q_x$). The connection form ω thus defined satisfies (iii) $R_a^*(\omega) = \text{ad}(a^{-1})\omega$ ($a \in G$), where $R_a^*(\omega)$ is the \dagger differential form induced by the

transformation R_a from the differential form ω . Conversely, given a 1-form ω with values in \mathfrak{g} that satisfies conditions (i) and (ii), we can define a connection in P by defining vectors X such that $\omega(X)=0$ as the horizontal, and its connection form coincides with ω . Thus giving a connection in P is equivalent to giving a connection form in P .

In particular, when a principal fiber bundle P is trivial, i.e., when $P = M \times G$, we can identify the tangent vector space $T_x(P)$ at a point $x = (p, g)$ of P with the direct sum of $T_p(M)$ and $T_g(G)$. If we set $Q_x = T_p(M)$, then Q defines a connection in $P = M \times G$. Such a connection is called **flat**. When a connection can always be expressed as above locally, it is called **locally flat**. Since each principal fiber bundle is locally a product fiber bundle, we see that locally there exists a connection. If the base space M is \dagger paracompact, we can show the existence of connections on P .

F. Extension and Restriction of Connections

When a principal fiber bundle $P = (P, \pi, M, G)$ has a \dagger reduced fiber bundle P' , we shall consider the relation between the connections of P and of P' . Let G' be a Lie subgroup of G and \mathfrak{g}' its Lie algebra. We shall denote by j both the injection of G' into G and also the injection of \mathfrak{g}' into \mathfrak{g} . If there exist a differentiable principal fiber bundle $P' = (P', \pi', M, G')$ and a differentiable embedding f of P' into P such that $\pi \circ f = \pi'$ and $f \circ R_a = R_{j(a)} \circ f$ ($a \in G'$) are satisfied, then (P', f) is said to be a reduced fiber bundle of P . Then we have $f_*(A_x^*) = j(A)_{f(x)}^*$ for each $A \in \mathfrak{g}'$ and $x \in P'$.

Suppose that a connection is given in P' ; we denote the horizontal space at the point x of P' by Q'_x . At the point $f(x)$ of P , we take $f_*(Q'_x)$ as the horizontal space and transform it by right translations of G . Thus we obtain a connection on P . Let ω' and ω be the corresponding connection forms. Then we have $j \circ \omega' = f^*(\omega)$ on P' . Conversely, suppose that we are given a connection in P with the connection form ω . If the induced form $f^*(\omega)$ on P' has values always in $j(\mathfrak{g}')$, we can write $f^*(\omega) = j \circ \omega'$, and ω' defines a connection in P' . In this case the connection in P is called an **extension** of the connection in P' , and the connection in P' is called the **restriction** of the connection in P .

G. Curvature Forms

Suppose that a principal fiber bundle $P = (P, \pi, M, G)$ has a connection. Let F be a finite-dimensional vector space and α be a differential form of degree k on P with values in F . We

define the **covariant differential** $D\alpha$ of α by

$$(D\alpha)(X_1, \dots, X_{k+1}) = (d\alpha)(hX_1, \dots, hX_{k+1}),$$

where the X_i are vector fields on P and h denotes the projection to the horizontal component. $D\alpha$ is a differential form of degree $k+1$ on P with values in F .

Let $\rho: G \rightarrow GL(F)$ be a \dagger representation of a Lie group G on F . A differential form α on P with values in F is called a **pseudotensorial form** of type ρ if α satisfies $R_a^*(\alpha) = \rho(a^{-1})\alpha$ ($a \in G$). In particular, if a pseudotensorial form α satisfies $\iota(A^*)\alpha = 0$ for any $A \in \mathfrak{g}$ (\rightarrow 105 Differentiable Manifolds Q), it is called a **tensorial form** of type ρ . For each representation ρ of G , we can construct an associated vector bundle E over M with fiber F . A tensorial form of type ρ is identified with a differential form on M with values in E . If α is a pseudotensorial form of type ρ , then $D\alpha$ is a tensorial form of type ρ .

For a connection form ω on P , the covariant differential $D\omega = \Omega$ of ω is called the **curvature form** of the connection. Since ω is a pseudotensorial form of type ad , Ω is a tensorial form of type ad . For the connection form we have the **structure equation** $d\omega = -[\omega, \omega] + \Omega$ [4, 6]. Let X and Y be vector fields on M , and let X^* and Y^* be their lifts, respectively. Then we have $\omega([X^*, Y^*]) = \Omega(X^*, Y^*)$, which shows that the curvature form Ω for X^*, Y^* gives the vertical component of $[X^*, Y^*]$.

It is known that the following three conditions for a connection are equivalent: (i) The connection is locally flat. (ii) The curvature form vanishes. (iii) The restricted holonomy group is trivial (i.e., the identity group).

The following two theorems are fundamental:

(1) Suppose that a connection is given in a principal fiber bundle $P = (P, \pi, M, G)$. Then the structure group of P can be reduced to the holonomy group [4, 6]. In fact, for $x \in P$, let $P(x)$ be the set of points y in P that can be connected to x by a piecewise horizontal curve in P . Then $P(x)$ gives a reduced fiber bundle of P , and the connection in P is an extension of a connection in $P(x)$ [4, 6].

(2) The Lie algebra of the holonomy group with a reference point x in P coincides with the vector subspace of \mathfrak{g} spanned by $\{\Omega_y(X, Y) \mid y \in P(x), X, Y \in T_y(P)\}$ [4, 6].

The curvature form Ω is used to express the \dagger characteristic classes of the bundle P [1, 2] (\rightarrow 56 Characteristic Classes).

In some cases, a connection in the principal fiber bundle induces a connection in an \dagger associated fiber bundle. In particular, when G is $GL(n, \mathbf{R})$ or $GL(n, \mathbf{C})$, we can define a connection in any associated vector bundle. The notion of connections in vector bundles can be

defined more algebraically (M. F. Atiyah, *Trans. Amer. Math. Soc.*, 85 (1957)) and can also be defined as a kind of differential operator on vector bundles [8].

II. Affine Connections

Let M be a differentiable manifold of dimension n and P be the \dagger bundle of tangent n -frames over M . Then P has the structure group $GL(n, \mathbf{R})$, and it is the principal bundle associated with the tangent vector bundle of M , which consists of all tangent vectors of M . A connection in the bundle of tangent n -frames is called an **affine connection** (or **linear connection**) on M . An affine connection on M defines (as well as the curvature form Ω) a new form Θ called the torsion form on P , which is given as follows: Let F be an n -dimensional vector space with a fixed basis $(\xi_1, \xi_2, \dots, \xi_n)$. Since the bundle of tangent n -frames P is the set of all bases (i.e., n -frames) (e_1, \dots, e_n) in $T_p(M)$ at each point p of M , every point $x = (e_1, \dots, e_n)$ of P is given as a mapping \bar{x} of F onto $T_p(M)$ ($p = \pi(x)$) defined by $\bar{x}: \xi_i \rightarrow e_i$. We define differential form θ of degree 1 with values in F on P by $\theta_x(X) = \bar{x}^{-1}(\pi_*(X))$ ($X \in T_x(P)$). θ is called a **canonical 1-form** of the bundle of tangent n -frames of the manifold M and has the following property: Any diffeomorphism φ of M onto itself induces a bundle automorphism $\tilde{\varphi}$ of P onto itself, and $\tilde{\varphi}$ preserves θ , that is, $\tilde{\varphi}^*(\theta) = \theta$. Conversely, we can show that any bundle automorphism of P that preserves θ is induced by a diffeomorphism of the base space M .

For an affine connection on M , we define the **torsion form** Θ by $\Theta = D\theta$. Θ is a differential form of degree 2 on P with values in F and satisfies $R_a\Theta = a^{-1} \cdot \Theta$ ($a \in GL(n, \mathbf{R})$). Furthermore, we have the **structure equation** for Θ , $d\theta = [\omega, \theta] + \Theta$ [2, 4, 6].

For each element ξ in F , there exists a unique horizontal vector field $B(\xi)$ on P such that $\theta(B(\xi)) = \xi$. $B(\xi)$ is called the **basic vector field** corresponding to ξ . At each point $x \in P$, $B(\xi_1)_x, \dots, B(\xi_n)_x$ form a basis of \mathcal{Q}_x . Let $\{A_1, \dots, A_m\}$ ($m = n^2$) be a basis of $\mathfrak{g} = \mathfrak{gl}(n, \mathbf{R})$. Then at each point $x \in P$, $\{(A^*)_1, \dots, (A^*)_m\}$, $B(\xi_1)_x, \dots, B(\xi_n)_x$ is a basis of the tangent vector space $T_x(P)$. Thus the bundle P of frames is a \dagger parallelizable manifold. The projection to M of any \dagger integral curve of a basic vector field is a geodesic, which is defined in Section I [4].

An affine connection on M gives a parallel displacement of the tangent vector space of M as follows: Let $C = p_t$ ($0 \leq t \leq 1$) be a curve in M and $C^* = x_t$ be a lift of C to P . The parallel displacement of the tangent n -frame x_0 at p_0

along the curve C is x_t , and the mapping $\bar{x}_t \circ \bar{x}_0^{-1}: T_{p_0}(M) \rightarrow T_{p_t}(M)$ is called the **parallel displacement** of the tangent space $T_{p_0}(M)$ onto $T_{p_t}(M)$ along C . It is easily seen that the mapping is independent of the choice of lifts.

I. Covariant Differentials

Let $C = \{p_t\}$ ($0 \leq t \leq 1$) be a differentiable curve in M . If we have a vector Y_t in $T_{p_t}(M)$ for each t and the correspondence $t \rightarrow Y_t$ is differentiable, then $\{Y_t\}$ is called a vector field along the curve C . For $\{Y_t\}$ we set

$$Y'_t = \lim_{h \rightarrow 0} (1/h)(\varphi_{t, t+h}^{-1}(Y_{t+h}) - Y_t),$$

where $\varphi_{t, t+h}$ is the parallel displacement of $T_{p_t}(M)$ onto $T_{p_{t+h}}(M)$ along the curve C . The vector field $\{Y'_t\}$ along C thus obtained is called the **covariant derivative** of $\{Y_t\}$. $\{Y_t\}$ is parallel along C ; that is, $Y_t = \varphi_{0, t}(Y_0)$ if and only if $Y'_t \equiv 0$. In particular, if the tangent vectors to a curve C are parallel along C itself, then C is said to be a **geodesic**.

Let X and Y be vector fields on a manifold M with an affine connection. The **covariant derivative** $\nabla_X Y$ of the vector field Y in the direction of the vector field X is defined as follows: Let p_0 be a point in M , $C = \{p_t\}$ ($-\varepsilon \leq t \leq \varepsilon$) be an integral curve of X through p_0 , and $\{\varphi_t\}$ be the parallel displacement along C . We set

$$(\nabla_X Y)_{p_0} = \lim_{t \rightarrow 0} (1/t)(\varphi_t^{-1}(Y_{p_t}) - Y_{p_0}).$$

Then $\nabla_X Y$ is also a vector field on M .

The mapping $(X, Y) \rightarrow \nabla_X Y$ satisfies the following three conditions: (i) $\nabla_X Y$ is linear with respect to X and Y ; (ii) $\nabla_{fX} Y = f \cdot \nabla_X Y$; and (iii) $\nabla_X(fY) = f \cdot \nabla_X Y + (Xf) \cdot Y$, where f is a differentiable function on M . Conversely, if a mapping satisfying conditions (i)–(iii) above is given, then there exists a unique affine connection on M whose covariant derivative coincides with the given mapping [4, 6].

Fix a vector field Y . Then the mapping $X \rightarrow \nabla_X Y$ defines a \dagger tensor field of type $(1, 1)$. This tensor field is called the **covariant differential** of Y and is denoted by ∇Y . Now fix a vector field X . Then the mapping $Y \rightarrow \nabla_X Y$ can be naturally extended to tensor fields of arbitrary type, and it commutes with the \dagger contraction of the tensors. For a tensor field K this is denoted by $K \rightarrow \nabla_X K$. Furthermore, the mapping $X \rightarrow \nabla_X K$ is called the **covariant differential** of K and is denoted by ∇K . We call $\nabla_X K$ the **covariant derivative** of K in the direction of X . A tensor field K is invariant under parallel displacements if and only if $\nabla K = 0$ (\rightarrow 417 Tensor Calculus).

J. Curvature Tensors and Torsion Tensors

For an affine connection on M , the **curvature tensor** R and the **torsion tensor** T are defined by

$$R(X, Y)(Z) = \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X, Y]}(Z),$$

$$T(X \cdot Y) = \nabla_X Y - \nabla_Y X - [X, Y],$$

where X, Y , and Z are vector fields on M , and R and T are tensors of types $(1, 3)$ and $(1, 2)$, respectively. Also, in terms of the curvature form Ω and the torsion form Θ on the bundle of tangent n -frames P over M , they can be defined by

$$R_p(X, Y)(Z) = \bar{x}^{-1} \cdot \Omega_x(X^*, Y^*) \cdot \bar{x}(Z),$$

$$T_p(X, Y) = \bar{x}^{-1}(\Theta_x(X^*, Y^*)),$$

where $\pi(x) = p$, $X, Y \in T_p(M)$, and X^*, Y^* are lifts of X, Y , respectively. The curvature tensor and the torsion tensor satisfy the relations $R(X, Y) = -R(Y, X)$, $T(X, Y) = -T(Y, X)$. Moreover, **Bianchi's identities** hold:

$$\begin{aligned} \mathfrak{S}(R(X, Y)(Z)) &= \mathfrak{S}(T(T(X, Y), Z) \\ &\quad + (\nabla_X T)(Y, Z)) \end{aligned}$$

and

$$\mathfrak{S}((\nabla_X R)(Y, Z) + R(T(X, Y), Z)) = 0,$$

where \mathfrak{S} denotes the sum of terms that are obtained by cyclic permutations of X, Y, Z [4]. For instance, in the case of a Riemannian connection (Section K), we have $T=0$, and Bianchi's identities reduce to

$$\begin{aligned} R(X, Y)(Z) + R(Y, Z)(X) + R(Z, X)(Y) &= 0, \\ (\nabla_X R)(Y, Z) + (\nabla_Y R)(Z, X) + (\nabla_Z R)(X, Y) &= 0. \end{aligned}$$

We now consider a system of coordinates (x^1, \dots, x^n) in an n -dimensional linear space $M = \mathbf{R}^n$. The vector fields (X_1, \dots, X_n) ($X_i = \partial/\partial x^i$) form a basis for vector fields on M . If we set $\nabla_{X_i} X_j = 0$, we get an affine connection on \mathbf{R}^n . For such a connection we have $R=0$, $T=0$, and any straight line in \mathbf{R}^n is a geodesic with respect to this connection. The connection is called the **canonical affine connection** on \mathbf{R}^n . An affine connection on a manifold M satisfies $R=0$ and $T=0$ if and only if the connection on M is locally isomorphic to the canonical affine connection of \mathbf{R}^n .

Let φ be a diffeomorphism of a manifold M with an affine connection onto itself. We call φ an **affine transformation** of M if the induced automorphism $\tilde{\varphi}$ on the bundle of tangent n -frames preserves the connection. In terms of covariant differentials, this condition is equivalent to the condition $\nabla_{\varphi(X)} \varphi(Y) = \varphi(\nabla_X Y)$ for any vector fields X, Y . An affine transformation of the canonical affine connection on \mathbf{R}^n

is an ordinary affine transformation. For an affine connection on a manifold M , the set of all affine transformations forms a Lie group and acts on M as a Lie transformation group [4, 5].

Let M be a manifold with an affine connection. M is called an affine **locally symmetric space** if $\nabla R=0$ and $\nabla T=0$ are satisfied. These conditions are satisfied if and only if at each point p of M , there exist a neighborhood U of p and an affine transformation φ of U such that $\varphi^2=1$ and p is the isolated fixed point of φ . If for each point p of M there exists an affine transformation of M such that $\varphi^2=1$ and p is an isolated fixed point, then M is called an **affine symmetric space**. A symmetric Riemannian space is a special case of this type (\rightarrow 413 Symmetric Spaces).

At each point p in a manifold M with an affine connection, we can choose local coordinates (x^1, \dots, x^n) such that $x^i(p)=0$ and the curve $x^i = a^i t$ ($-\delta < t < \delta$) is a geodesic for each (a^1, \dots, a^n) with $\sum (a^i)^2 = 1$. Such local coordinates are called **geodesic coordinates** at p [4]. With respect to geodesic coordinates, we have $(\nabla_{X_i}(X_j))_p = 0$ ($\partial/\partial x^i = X_i$).

K. Riemannian Connections

When a Riemannian metric g (\rightarrow 364 Riemannian Manifolds) is given on a manifold M , it defines a metric on the tangent space $T_p(M)$ at each point p of M , and we can take orthonormal bases in $T_p(M)$. The set P' of all orthonormal bases of tangent spaces is a subset of the bundle P of tangent n -frames of M and forms a subbundle of P ; its structure group is the orthogonal group $O(n)$, and P' gives a reduction of the bundle of tangent n -frames. Conversely, when a reduction of the bundle of frames to $O(n)$ is given, we can define a Riemannian metric on M such that the reduced bundle consists of all orthonormal frames.

For a Riemannian metric g on M , there exists a unique affine connection on M such that (i) $\nabla g=0$, and (ii) the torsion tensor T vanishes [4]. This connection is called the **Riemannian connection** corresponding to g . The first condition is equivalent to the invariance of the Riemannian metric g under parallel displacement. Thus the affine connection transforms orthonormal bases on M to orthonormal bases and induces a connection in the bundle P' . It is known that the restricted holonomy group of any Riemannian connection is a closed subgroup of $O(n)$ [4]. An affine connection on a manifold M is called a **metric**

connection if it preserves a Riemannian metric g on M , i.e., if it satisfies the condition (i).

L. Representations in Local Coordinates

(1) Let (x^1, \dots, x^n) be a local coordinate system in a manifold M and consider vector fields $X_i = \partial/\partial x_i$. For an affine connection on M , the covariant derivative can be expressed as

$$\nabla_{X_j}(X_k) = \sum_i \Gamma_{jk}^i X_i.$$

The Γ_{jk}^i are called **coefficients of the affine connection** with respect to the local coordinate system (x^1, \dots, x^n) . We denote by $\bar{\Gamma}_{jk}^i$ the coefficients of connection with respect to another local coordinate system (y^1, \dots, y^n) . Then on the intersection of their coordinate neighborhoods, we have

$$\bar{\Gamma}_{jk}^i = \sum_{\alpha, \beta, \gamma} \frac{\partial y^i}{\partial x^\alpha} \left(\frac{\partial x^\beta}{\partial y^j} \frac{\partial x^\gamma}{\partial y^k} \Gamma_{\beta\gamma}^\alpha + \frac{\partial^2 x^\alpha}{\partial y^j \partial y^k} \right).$$

Conversely, if the Γ_{jk}^i are given in each local coordinate system of M and satisfy this relation on each intersection of their coordinate neighborhoods, then there exists a unique affine connection such that the coefficients of the connection are given by Γ_{jk}^i .

(2) The **coefficients of the Riemannian connection** corresponding to a Riemannian metric $g = \sum g_{ij} dx^i dx^j$ on a manifold M are given by

$$\Gamma_{jk}^i = \frac{1}{2} \sum_l g^{il} \left\{ \frac{\partial g_{jl}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l} \right\}$$

and are called the **Christoffel symbols**.

(3) With respect to each local coordinate system (x^1, \dots, x^n) , we express the **torsion tensor** T and the **curvature tensor** R of an affine connection by

$$T = \sum_{ijk} T_{jk}^i dx^j \otimes dx^k \otimes X_i,$$

$$R = \sum_{ijkl} R_{jkl}^i dx^j \otimes dx^k \otimes dx^l \otimes X_i.$$

The components T_{jk}^i and R_{jkl}^i are given by

$$\begin{aligned} T_{jk}^i &= \Gamma_{jk}^i - \Gamma_{kj}^i, \\ R_{jkl}^i &= (\partial \Gamma_{lj}^i / \partial x^k - \partial \Gamma_{kj}^i / \partial x^l) \\ &\quad + \sum_m (\Gamma_{lj}^m \Gamma_{km}^i - \Gamma_{kj}^m \Gamma_{lm}^i). \end{aligned}$$

(4) Let $K = (K_{j_1 \dots j_s}^{i_1 \dots i_r})$ be a tensor field of type (r, s) . Then the **covariant differential** $\nabla K = (K_{j_1 \dots j_s, k}^{i_1 \dots i_r})$ is given by

$$\begin{aligned} K_{j_1 \dots j_s, k}^{i_1 \dots i_r} &= \partial K_{j_1 \dots j_s}^{i_1 \dots i_r} / \partial x^k \\ &\quad + \sum_{\alpha=1}^r \left(\sum_l \Gamma_{kl}^\alpha K_{j_1 \dots j_s}^{i_1 \dots i_{\alpha-1} \alpha i_{\alpha+1} \dots i_r} \right) \\ &\quad - \sum_{\beta=1}^s \left(\sum_m \Gamma_{kj}^m K_{j_1 \dots j_{\beta-1} m j_{\beta+1} \dots j_s}^{i_1 \dots i_r} \right). \end{aligned}$$

(5) A curve $x^i = x^i(t)$ is a **geodesic** if and only if

$$\frac{d^2 x^i}{dt^2} + \sum_{j,k} \Gamma_{jk}^i \frac{dx^j}{dt} \frac{dx^k}{dt} = 0, \quad i = 1, 2, \dots, n$$

(\rightarrow 178 Geodesics, 417 Tensor Calculus).

M. Cartan Connections

Let M be a differentiable manifold of dimension n . Consider a homogeneous space $F = G/G'$ of the same dimension n , where G is a Lie group and G' is a closed subgroup of G (\rightarrow 199 Homogeneous Spaces). Let $B = (B, M, F, G)$ be a fiber bundle over M with fiber F and structure group G , and $P = (P, M, G)$ be the principal fiber bundle associated with B . Suppose that there exists a cross section f over M to B . Then the structure group of P can be reduced to G' . We denote this reduced fiber bundle by $P' = (P', M, G')$ and the injection of P' into P by j (\rightarrow 147 Fiber Bundles).

Suppose that a connection is given in P . Its connection form ω is a differential form of degree 1 on P with values in \mathfrak{g} , and the induced form $\omega' = j^*(\omega)$ is also a differential form of degree 1 on P' with values in \mathfrak{g} . We call the connection in P a **Cartan connection** on M with the fiber $F = G/G'$ if at each point x of P' , ω'_x gives an isomorphism of $T_x(P')$ onto \mathfrak{g} as linear spaces. Such a connection in P is equivalently defined as a 1-form ω' on P' with values in \mathfrak{g} satisfying the following three conditions: (i) $\omega'(A^*) = A$ ($A \in \mathfrak{g}'$ (Lie algebra of G')); (ii) $R_\#^*(\omega') = \text{ad}(a^{-1})\omega'$ ($a \in G'$); and (iii) ω'_x gives an isomorphism of $T_x(P')$ onto \mathfrak{g} at each point $x \in P'$. For such ω' , we can take a connection form ω in P such that $\omega' = j^*(\omega)$; ω defines a Cartan connection.

N. Soudures

A cross section f over M to B gives a vector bundle $T'(B)$ on M defined as follows: For each point p of M , the projection $B \rightarrow M$ defines a mapping $T_{f(p)}(B) \rightarrow T_p(M)$. The kernel of this mapping is denoted by $V_{f(p)}(B)$. Then $T'(B) = \bigcup_p V_{f(p)}(B)$ forms a vector bundle over M , and the dimension of its fibers is equal to $n = \dim F$.

A Cartan connection in P gives a bundle isomorphism between $T'(B)$ and the tangent vector bundle $T(M)$ of M as follows: Let x be an arbitrary point in P' , and put $p = \pi(x)$. The projection $\pi: P' \rightarrow M$ induces an isomorphism of $T_x(P')/V_x(P')$ onto $T_p(M)$. On the other hand, ω'_x gives an isomorphism of $T_x(P')/V_x(P')$ onto $\mathfrak{g}/\mathfrak{g}'$. As a point in P' , x gives a mapping of F

$= G/G'$ onto the fiber in B over p and sends the point $\{G'\}$ in F to $f(p)$. By this mapping, $T_0(F) = \mathfrak{g}/\mathfrak{g}'$ is mapped isomorphically onto $V_{f(p)}(B)$. Combining these isomorphisms, we get an isomorphism between $T_p(M)$ and $V_{f(p)}(B)$ that is independent of the choice of $x \in P'$ over p . The set of such isomorphisms for $p \in M$ defines a bundle isomorphism of $T(M)$ and $T'(B)$. If a fiber bundle B over M has an isomorphism such as above through a cross section, then B is said to have a **soudure** [3].

Conversely, if a fiber bundle B over M has a soudure with respect to a cross section f , then there exists a Cartan connection in P such that the soudure given by the connection coincides with the original one [3]. There are many Cartan connections in P with the given soudure. However, when $F = G/G'$ is a symmetric space of compact type such that G is noncompact and contains the identity component of the group of isometries, we can determine uniquely the so-called **normal** Cartan connection among the Cartan connections which gives rise to the given soudure [10].

For the tangent vector bundle $T(M)$ of M , the fiber F is an n -dimensional linear space and can be expressed as $F = G/G'$, where G is the \dagger affine transformation group of F and $G' = GL(n, \mathbf{R})$. Then $T(M)$ has the 0-section over M , and there exists a natural soudure. Furthermore, an affine connection on M canonically induces a Cartan connection on M with the fiber $F = G/G'$ [3].

For a Cartan connection on M , we can introduce the notion of **development** of a curve in M into the fiber and also the notion of **completeness** [3].

O. Projective Connections

Let $F_1 = G_1/G'_1$ and $F_2 = G_2/G'_2$ be homogeneous spaces with $\dim F_1 = \dim F_2 = n$. Suppose that G_1 is a Lie subgroup of G_2 and G'_1 is contained in G'_2 by the injection. Then we have a canonical injection $F_1 \rightarrow F_2$ (F_1 is an open subset of F_2 by the assumption).

Suppose that a fiber bundle B_1 with fiber F_1 over M has a cross section f_1 . Using f_1 , we can construct a bundle B_2 with fiber F_2 over M which also has a cross section f_2 . The principal bundle of B_2 is given by extending the structure groups from the principal bundle of B_1 . We can show that if B_1 has a soudure with respect to the cross section f_1 , then B_2 also has a soudure with respect to the cross section f_2 . A Cartan connection in the principal fiber bundle P_1 associated with B_1 that is compatible with a given soudure on B_1 induces a Cartan connection in the principal fiber bundle P_2 associated with B_2 , which then

induces a soudure on B_2 . The latter is called a Cartan connection **induced** from the former.

Let F_1 be an n -dimensional linear space and F_2 be the real \dagger projective space of dimension n . Then the affine transformation group of F_1 can be embedded into the projective transformation group of F_2 . Thus the tangent vector bundle of a manifold M induces a fiber bundle over M with the n -dimensional projective space as its fiber. A Cartan connection in this fiber bundle is called a **projective connection** on M . By the argument in this section, we see that every affine connection on M induces a projective connection on M .

Given two affine connections on M , we denote by ∇ and ∇' their corresponding covariant differentials. The two affine connections on M induce the same projective connection on M if and only if there exists a differential form φ of degree 1 on M such that $\nabla'_x Y - \nabla_x Y = \varphi(X)Y + \varphi(Y)X$ for any vector fields X, Y [7]. If a diffeomorphism φ of M preserves the projective connection induced by an affine connection in M , then φ maps geodesics of M into geodesics.

P. Conformal Connections

Let F_1 be an n -dimensional Euclidean space and F_2 an n -dimensional sphere (a \dagger conformal space). We can embed the group of \dagger isometries of F_1 canonically into the group of \dagger conformal transformations of F_2 . A Riemannian metric of the tangent vector bundle of a manifold M of dimension n gives a fiber bundle over M with fiber F_2 . A Cartan connection in this fiber bundle is called a **conformal connection** on M ; a Riemannian connection on M induces a conformal connection on M .

Two Riemannian metrics g_1, g_2 on M induce the same conformal connection on M if and only if there exists a positive function f on M such that $g_2 = fg_1$. Thus for a Riemannian manifold M with metric tensor g_1 , a diffeomorphism φ of M such that $\varphi^*(g) = fg_1$ leaves invariant the conformal connection induced by g_1 . Such a φ is called a **conformal transformation** of M with respect to the given Riemannian metric g_1 .

For a Riemannian manifold M with metric tensor g , we define **Weyl's conformal curvature tensor** W by

$$W_{jkl}^i = R_{jkl}^i + \frac{1}{n-2}(R_{jk}\delta_l^i - R_{jl}\delta_k^i + g_{jk}R_l^i - g_{jl}R_k^i) - \frac{R}{(n-1)(n-2)}(g_{jk}\delta_l^i - g_{jl}\delta_k^i),$$

where the R_{jkl}^i and R_{jk} are components of the curvature tensor and Ricci tensor, respectively,

and R is the scalar curvature (\rightarrow 364 Riemannian Manifolds, 417 Tensor Calculus). When $\dim M \geq 3$, the conformal connection induced by g on M is locally flat if and only if the conformal curvature tensor vanishes [7].

Q. Yang-Mills G -Connection

Let $P = (P, \pi, M, G)$ be a differentiable principal fiber bundle over a compact oriented Riemannian manifold M with group G . Then fiber bundles $G_P = P \times_c G$ and $\mathfrak{g}_P = P \times_{\text{Ad}} \mathfrak{g}$ associated with P are induced naturally from the group conjugation $c: G \rightarrow \text{Aut}(G)$ and the adjoint representation $\text{Ad}: G \rightarrow \text{Aut}(\mathfrak{g})$, respectively. A (local) section of G_P is called a (local) **gauge transformation** of P . The set of all global gauge transformations, which is denoted by \mathcal{G}_P , has a group structure.

A locally faithful representation ρ of G to an n -dimensional complex vector space F with a fixed basis (ξ_1, \dots, ξ_n) defines a differentiable complex vector bundle $E = P \times_\rho F$ associated with P . Every point x of P is identified with a linear mapping \bar{x} of F onto the fiber $E_{\pi(x)}$ defined by $\bar{x}: \xi_i \rightarrow e_i$, where e_i denotes the equivalence class of $\{x, \xi_i\} \in P \times F$, $1 \leq i \leq n$.

In a manner similar to the case of an affine connection, a connection in P with connection form ω gives a notion of parallel displacement of E as follows: Let $c = p_t$ ($0 \leq t \leq 1$) be a curve in M and $c^* = x_t$ be a lift of c to P . The mapping $\bar{x}_t \circ \bar{x}_0^{-1}: E_{p_0} \rightarrow E_{p_t}$ is called the parallel displacement of E_{p_0} onto E_{p_t} along c .

Let X be a vector field on M and φ be a differentiable section of E . The covariant derivative $\nabla_X \varphi$ of φ in the direction of X is defined as follows: Let p_0 be a point of M , $c = p_t$ ($-\varepsilon \leq t \leq \varepsilon$) be an integral curve of X through p_0 and $c^* = x_t$ be a lift of c to P . We set

$$(\nabla_X \varphi)_{p_0} = \lim_{t \rightarrow 0} \frac{1}{t} (\bar{x}_t \circ \bar{x}_0^{-1}(\varphi_{p_t}) - \varphi_{p_0}).$$

Then ∇_X is also a differentiable section of E .

The mapping $(X, \varphi) \rightarrow \nabla_X \varphi$ satisfies the following conditions: (i) $\nabla_X \varphi$ is linear with respect to X and φ ; (ii) $\nabla_{fX} \varphi = f \nabla_X \varphi$; and (iii) $\nabla_X (f\varphi) = (Xf)\varphi + f \nabla_X \varphi$, where f is a differentiable function on M . From these conditions it is seen that the mapping $X \rightarrow \nabla_X \varphi$ for a fixed section φ of E defines a differential form of degree one with values in E , denoted by $\nabla \varphi$. ∇ is called a G -connection on E (induced from the connection in P). A linear operator $d^\nabla: \Gamma(\Lambda^p \otimes E) \rightarrow \Gamma(\Lambda^{p+1} \otimes E)$, defined by $d^\nabla(\alpha \otimes \varphi) = d\alpha \otimes \varphi + (-1)^p \alpha \wedge \nabla \varphi$ for a differential form α of degree p and a differentiable section φ of E is called a covariant exterior differentiation. Here $\Gamma(\Lambda^p \otimes E)$ denotes the set of all differential forms of degree p with values in E .

The curvature form R^∇ of ∇ is defined by $R^\nabla(X, Y)\varphi = \nabla_X(\nabla_Y \varphi) - \nabla_Y(\nabla_X \varphi) - \nabla_{[X, Y]}\varphi$, where X and Y are vector fields on M and φ is a differentiable section of E . In terms of the curvature form Ω of the connection in P , it can be defined also by

$$R_P^\nabla(X, Y)\varphi = \bar{x} \circ \Omega_X(X^*, Y^*) \cdot \bar{x}^{-1}(\varphi),$$

where X and Y are in $T_p M$ and $x \in \pi^{-1}(p)$ and X^*, Y^* are lifts of X, Y to P , respectively, and φ is an element of E_p . Since Ω is a tensorial form of type ad and the differential ρ_* of ρ induces a faithful representation of \mathfrak{g} to $\mathfrak{gl}(F)$, R^∇ can be regarded as a differential form of degree two with values in the bundle $\mathfrak{g}_P = P \times_{\text{Ad}} \mathfrak{g}$.

The curvature form R^∇ satisfies Bianchi's identity: $d^\nabla R^\nabla = 0$, where d^∇ is a covariant exterior differentiation with respect to the G -connection on \mathfrak{g}_P canonically induced from the connection in P .

We denote by \mathcal{C} the set of all G -connections on E . Now let G be a compact semisimple Lie group. A functional $\mathcal{S}: \mathcal{C} \rightarrow \mathbf{R}$ defined by $\nabla \rightarrow \mathcal{S}(\nabla) = -\frac{1}{2} \int_M \text{tr}(R^\nabla \wedge * R^\nabla)$ is called the **action integral (Yang-Mills functional)** of ∇ , where $*$ is Hodge's star operator, given by the fixed orientation of M .

The group \mathcal{G}_P acts on \mathcal{C} by $\nabla \rightarrow f^{-1} \circ \nabla \circ f$ for $f \in \mathcal{G}_P$ and then the curvature form is transformed by this action as $R^\nabla \rightarrow \text{Ad}(f^{-1})R^\nabla$. Thus \mathcal{S} is \mathcal{G}_P -invariant.

A connection ∇ is called a **Yang-Mills G -connection** if ∇ is a critical point of \mathcal{S} . The Euler-Lagrange equation of \mathcal{S} is given by $d^{\nabla*} R^\nabla = 0$ by the aid of the formal adjoint operator $d^{\nabla*}$ of d^∇ . This equation, called the **Yang-Mills equation**, is a system of nonlinear second-order elliptic partial differential equations.

When M is a 4-dimensional vector space \mathbf{R}^4 with Minkowskian metric and G is the Abelian group $U(1)$, the Yang-Mills equation coincides with Maxwell's equations for an electromagnetic field. Thus the Yang-Mills equation for a non-Abelian group G is a natural extension of Maxwell's equations. In fact, the theory of Yang-Mills connections has its origin in the field theory of physics [11].

In the case of $\dim M = 4$, special Yang-Mills G -connections occur. A G -connection ∇ satisfying the condition $*R^\nabla = R^\nabla$ (resp. $*R^\nabla = -R^\nabla$) as a differential form of degree two is called a **self-dual** (resp. **anti-self-dual**) G -connection. From Bianchi's identity and the expression of $d^{\nabla*}: d^{\nabla*} = -* \circ d^\nabla \circ *$, it follows that every self-dual (anti-self-dual) G -connection gives a solution to the Yang-Mills equation. Since the first Pontryagin number $p_1(E)$ is given by $p_1(E) = -(1/4\pi^2) \int_M \text{tr}(R^\nabla \wedge R^\nabla)$ by virtue of the Chern-Weil theorem, the

action integral satisfies $\mathcal{S}(\nabla) \geq 2\pi^2 |p_1(E)|$ for every ∇ in \mathcal{C} and the equality holds if and only if ∇ is self-dual or anti-self-dual.

Explicit forms have been obtained for (anti-) self-dual connections over the 4-sphere S^4 by many interesting methods [12–15]. And it has been shown that moduli space of self-dual G -connections (i.e., the set of all solutions to the Yang-Mills equation modulo \mathcal{G}_P) over S^4 has the structure of a Hausdorff manifold of dimension $p_1(\mathfrak{g}_P) - \dim G$ for every principal bundle P with group G [16, 17]. It is not yet known whether there exists a Yang-Mills G -connection over S^4 whose holonomy group is an open subgroup of G and which is neither self-dual nor anti-self-dual [18]. The following is one of the few known facts concerning the properties of Yang-Mills G -connections: If a Yang-Mills G -connection, $G = SU(2)$, $SU(3)$, or $O(3)$, over S^4 is weakly stable, i.e., if the second variation of \mathcal{S} is positive semidefinite, then it is self-dual or anti-self-dual [19].

References

- [1] H. Cartan, *Notions d'algèbre différentielle; application aux groupes de Lie et aux variétés où opère un groupe de Lie*, Colloque de Topologie, Brussels, 1950, 15–27.
- [2] S. S. Chern, *Topics in differential geometry*, Lecture notes, Institute for Advanced Study, Princeton, 1951.
- [3] C. Ehresmann, *Les connexions infinitésimales dans un espace fibré différentiable*, Colloque de Topologie, Brussels, 1950, 29–55.
- [4] S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, Interscience, I, 1963; II, 1969.
- [5] A. Lichnerowicz, *Théorie globale des connexions et des groupes d'holonomie*, Cremona, 1955.
- [6] K. Nomizu, *Lie groups and differential geometry*, Publ. Math. Soc. Japan, 1956.
- [7] K. Yano and S. Bochner, *Curvature and Betti numbers*, Ann. Math. Studies, Princeton Univ. Press, 1953.
- [8] R. S. Palais, *Seminar on the Atiyah-Singer index theorem*, Ann. Math. Studies, Princeton Univ. Press, 1965.
- [9] S. S. Chern, *Differentiable manifolds*, Lecture notes, Univ. of Chicago, 1959.
- [10] N. Tanaka, *On the equivalence problem associated with a certain class of homogeneous spaces*, J. Math. Soc. Japan, 17 (1965), 103–139.
- [11] C. N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Phys. Rev., 96 (1954), 191–195.
- [12] A. A. Belavin, A. M. Polyakov, A. S.

Schwartz, and Yu. Tyupkin, *Pseudo-particle solutions of the Yang-Mills equations*, Phys. Lett., 59B, 1 (1975), 85–87.

[13] A. A. Belavin and V. E. Zakharov, *Yang-Mills equations as inverse scattering problem*, Phys. Lett. 73B, (1978), 53–57.

[14] R. Jackiw, C. Nohl, and C. Rebbi, *Conformal properties of pseudo particle configurations*, Phys. Rev. D15, 6 (1977), 1642–1646.

[15] V. G. Manin and Yu. I. Manin, *A description of instantons*, Comm. Math. Phys., 63 (1978), 177–192.

[16] M. F. Atiyah, N. J. Hitchin, and I. M. Singer, *Self-duality in four-dimensional Riemannian geometry*, Proc. Royal Soc. London A, 362 (1978), 425–461.

[17] M. Itoh, *Moduli of anti-self-dual connections on Kähler manifolds*, Proc. Japan Acad., 57A (1981), 176–180.

[18] M. F. Atiyah and J. D. S. Jones, *Topological aspects of Yang-Mills theory*, Comm. Math. Phys., 61 (1978) 97–118.

[19] J. P. Bourguignon and H. B. Lawson, *Stability and isolation phenomena for Yang-Mills fields*, Comm. Math. Phys., 79 (1981), 189–230.

81 (I.11) Constructive Ordinal Numbers

A. General Remarks

To extend the theory of \dagger recursive functions to transfinite ordinal numbers, A. Church and S. C. Kleene [1] considered the set of effectively accessible ordinal numbers and defined the concept of constructive ordinal numbers as explained later in this article. Their work became the basis of fruitful research by Kleene, W. Markwald, C. Spector, and others [2–5]. A constructive ordinal number was originally introduced as an “expression” in a \dagger formal system utilizing the λ -notation. Since such a system is “effective,” we can arithmetize it utilizing \dagger Gödel numbers and assume from the outset that each ordinal number is representable by a natural number. The notations, terminology, and theorems mentioned in this article are mainly those for constructive ordinal numbers of the \dagger second number class.

B. Definition and Fundamental Properties

We call a set of natural numbers satisfying conditions (I) and (II) a **system of notations for ordinal numbers**, and an ordinal number a

constructive ordinal number when it is representable by a natural number belonging to such a system of notations: (I) No natural number represents two distinct ordinal numbers. (II) There are three \dagger partial recursive functions $K(x)$, $P(x)$, and $Q(x, n)$ defined as follows: (i) for any natural number x representing X , $K(x)$ takes the value 0, 1, or 2 according as X is zero, an \dagger isolated ordinal number, or a \dagger limit ordinal number, respectively; (ii) when X is the ordinal number \dagger immediately after an ordinal number Y , $P(x)$ represents Y for any natural number x representing X ; (iii) when X is a limit ordinal number, for any natural number x representing X there exists an increasing sequence $\{Y_n\}$ of ordinal numbers such that $X = \lim_n Y_n$ and $Q(x, n)$ represents Y_n for each natural number n .

The system called S_3 by Kleene is the most useful and convenient among systems of notations for ordinal numbers. Let n_0 be a \dagger primitive recursive function of the variable n defined by $0_0 = 1$, $(n+1)_0 = 2^{n_0}$. The fundamental notion $a \in O$ and relation $a <_O b$ of the system S_3 are introduced by the following inductive definition: (1) $1 \in O$; (2) if $y \in O$, then $2^y \in O$ and $y <_O 2^y$; (3) if a sequence $\{y_n\}$ of natural numbers has the property that for each n , $y_n \in O$ and $y_n <_O y_{n+1}$, and if y is a Gödel number that defines y_n recursively as a function of n_0 (i.e., $y_n \cong \{y\}(n_0)$) (\rightarrow 356 Recursive Functions), then $3 \cdot 5^y \in O$, and for each n , $y_n <_O 3 \cdot 5^y$; (4) if $x, y, z \in O$, $x <_O y$, and $y <_O z$, then $x <_O z$; (5) $a \in O$, $a <_O b$ hold only when they follow from (1)–(4).

Now, for brevity, we write $a \leq_O b$ for $(a <_O b) \vee (a = b)$. The following propositions hold for S_3 : (1) If $a <_O b$, then $b \neq 1$; (2) If $a <_O b$, then $a, b \in O$; (3) If $a <_O 2^y$, then $a \leq_O y$; (4) If $a <_O 3 \cdot 5^y$, then there is a natural number n such that $a \leq_O y_n$, where $y_n = \{y\}(n_0)$; (5) If $a \in O$, then $1 \leq_O a$; (6) If $a \in O$, then for any \dagger number-theoretic function α such that $\alpha(0) = a$ and $\forall n(\alpha(n) \neq 1 \rightarrow \alpha(n+1) <_O \alpha(n))$, there is a k such that $\alpha(k) = 1$; (7) For each a , $\neg(a <_O a)$; (8) If $c \in O$, $a \leq_O c$, and $b \leq_O c$, then $a <_O b$ or $a = b$ or $b <_O a$.

Each member a of O represents an ordinal number $|a|$ as follows: $|1| = 0$; $|2^y| = |y| + 1$ for $y \in O$; $|3 \cdot 5^y| = \lim_n |y_n|$ for $3 \cdot 5^y \in O$, where $y_n = \{y\}(n_0)$. Let b be a member of O . Then $|a| < |b|$ when $a <_O b$; and conversely, for each $\alpha < |b|$, there is a number a such that $|a| = \alpha$ and $a <_O b$. Hence the set $\{a | a <_O b\}$ is a \dagger well-ordered set with respect to $<_O$, and its \dagger order type is $|b|$. The least number ξ greater than $|a|$ for every member a of O is the least ordinal number that is not constructive. It is denoted by ω_1^{CK} (Church and Kleene denoted it by ω_1). There is a subsystem of S_3 that is well ordered with respect to \leq_O and contains a

unique notation for each constructive ordinal number α . For such a subsystem we can take a Π_1^1 set K such that K is \dagger recursive in O (S. Feferman and Spector, R. O. Gandy) (\rightarrow 356 Recursive Functions H).

C. Constructive Ordinals and the Kleene Hierarchy

Let $R(x, y)$ be a \dagger predicate on natural numbers. We write $x \leq_R y$ for any natural numbers x, y for which $R(x, y)$ holds. We consider only the case where \leq_R is a \dagger linear ordering on the set $D_R = \{x | \exists y(R(x, y) \vee R(y, x))\}$. If D_R is a well-ordered set with respect to \leq_R , we denote its \dagger order type by $|R|$. (1) For each (constructive) ordinal number $\alpha < \omega_1^{\text{CK}}$, there is a \dagger general recursive (more strictly, \dagger primitive recursive) predicate R such that $|R| = \alpha$ (Markwald, Spector, Kleene [3]). (2) Conversely, if R is a \dagger hyperarithmetical predicate (e.g., R is general recursive), then $|R| < \omega_1^{\text{CK}}$ (Markwald, Spector). The following theorems are the most fruitful ones in the theory of constructive ordinal numbers, and they fully support the validity of Kleene's idea of \dagger analytic hierarchy. (3) The set O is Π_1^1 (\rightarrow 356 Recursive Functions H), and so is the predicate $a <_O b$. Namely, for O , there is a primitive recursive predicate $R(a, x, \alpha)$ such that

$$a \in O \Leftrightarrow \forall x \exists \alpha R(a, x, \alpha)$$

(Kleene [3]). (4) For each ordinal number $\alpha < \omega_1^{\text{CK}}$, the set $\{a | \alpha \geq |a|\}$ is a hyperarithmetical set (Spector). (5) O is a \dagger complete set for Π_1^1 . That is, for any Π_1^1 set E , there is a primitive recursive function φ such that $a \in E \Leftrightarrow \varphi(a) \in O$ (Kleene [3]). Accordingly, O is not a Σ_1^1 set (\rightarrow 356 Recursive Functions H).

D. Relativization and Extension

Given a (number-theoretic) predicate Q of one variable (or a set of natural numbers), we can \dagger relativize to Q the notion of constructive ordinal numbers. The least ordinal that is not constructive relative to Q is denoted by ω_1^Q . The relativization to Q of the fundamental notion $a \in O$ and relation $a <_O b$ of the system S_3 of notations are denoted by $a \in O^Q$ and $a <_{O^Q} b$, respectively. Then we can relativize the results of the preceding paragraphs to Q . For example, as the relativization of (3), we have the following: There is a predicate $R^Q(a, x, \alpha)$ which is primitive recursive \dagger uniformly in Q such that

$$a \in O^Q \Leftrightarrow \forall x \exists \alpha R^Q(a, x, \alpha).$$

When Q is hyperarithmetical, we have no generalization of the constructive ordinal numbers

Constructive Ordinal Numbers

by relativizing them to Q , that is, $\omega_1^Q = \omega_1^{CK}$ holds (Spector). Now by relativizing to O the concept of constructive ordinal numbers, we obtain a proper extension of it ($\omega_1^{CK} < \omega_1^O$), and then, performing such extensions successively, we have a (transfinite) sequence O, O^O, O^{OO}, \dots . On the other hand, we can extend the constructive ordinal numbers to those corresponding to any like higher number class, beyond the second, in which partial recursive functions are used at limit levels to provide an "accessibility" mapping from previously defined number classes. There are several extensions done by Church and Kleene, H. C. Wang, D. L. Kreider and H. Rogers, Jr., H. Putnam, W. Richter, A. Kino and G. Takeuti, and others. Richter [9] has shown that these two ways of extending the constructive ordinals are equivalent, provided the sets of notations for the higher number classes satisfy certain natural conditions. Specifically, the ordinals of the Addison and Kleene [6] constructive third number class are exactly the ordinals less than ω_1^O and the set O_{2^O} of notations for those is recursively isomorphic to O^O (Richter).

E. Constructive Ordinals in Higher Number Classes

Putnum [8] has defined a system C of notations for constructive ordinals of the Cantor higher number classes, improving that of Kreider and Rogers [7], as follows: N_α is the set of notations for the ordinal number α ; if for some α , $x \in N_\alpha$, let $|x| = \mu\xi [x \in N_\xi]$, which is called a **hyperconstructive ordinal**; and let $C_\alpha = \bigcup \{N_\xi \mid \xi < \alpha\}$. Define i to be a Gödel number of the \dagger identity function, and n to be an index in C_α if $3^a \cdot 5^i \in C_\alpha$ for some t . There are four cases:

Case 1. $\alpha = 0$. Then $N_\alpha = \{1\}$.

Case 2. $\alpha = \beta + 1$, where N_β is already defined. Then $N_\alpha = \{2^x \mid x \in N_\beta\}$.

Case 3. α is a limit ordinal such that N_ξ is already defined for all $\xi < \alpha$, and there exists an ordinal $\beta < \alpha$ such that $3^a \cdot 5^i \in N_\beta$ for some a , and a partial recursive function f giving an order-preserving \dagger cofinal mapping (o.p.c.m.) from C_β into C_α . Then N_α is taken to be the set of all numbers $3^a \cdot 5^n$ such that $3^a \cdot 5^i \in N_\beta$ (where β is any ordinal with the above property) and $\{n\}$ is an o.p.c.m. from C_β into C_α .

Case 4. α is a limit ordinal such that N_ξ is already defined for all $\xi < \alpha$, there is no β as in Case 3, but there is a number $a \in C_\alpha$ which is not an index in C_α . Let β_0 be the least β for which such a 's belong to N_β . Then $N_\alpha = \{3^a \cdot 5^n \mid a \in N_{\beta_0} \text{ and } \{n\} \text{ is an o.p.c.m. from } C_\alpha \text{ into } C_\alpha\}$.

We have that $x \in N_\alpha$ only as required by Cases 1–4. Finally, let ω_C be the least ordinal number which is not hyperconstructive, and let $C = C_{\omega_C}$.

Extending Kleene's H_y for $y \in O$ (\rightarrow 356 Recursive Functions H), Putnum has defined a hierarchy H_x for $x \in C$, and shown that hyperconstructive ordinal numbers are uniqueness ordinals, i.e., for $x, y \in C$, if $|x| = |y|$, then H_x and H_y are of the same \dagger degree of (recursive) unsolvability. E_1 is the type-2 object (\rightarrow 356 Recursive Functions F) introduced by T. Tugué such that $E_1(\alpha) = 0$ if $\forall \beta \exists x [\alpha(\beta(x)) = 0]$; otherwise, $E_1(\alpha) = 1$. It is known that $\omega_C = \omega^{E_1}$, where ω^{E_1} is the least ordinal which is not the order type of any well-ordering recursive in E_1 (Richter [10]).

References

- [1] A. Church and S. C. Kleene, Formal definitions in the theory of ordinal numbers, *Fund. Math.*, 28 (1937), 11–21.
- [2] S. C. Kleene, On notation for ordinal numbers, *J. Symbolic Logic*, 3 (1938), 150–155.
- [3] S. C. Kleene, On the forms of the predicates in the theory of constructive ordinals II, *Amer. J. Math.*, 77 (1955), 405–428.
- [4] W. Markwald, Zur Theorie der konstruktiven Wohlordnungen, *Math. Ann.*, 127 (1954), 135–149.
- [5] C. Spector, Recursive well-orderings, *J. Symbolic Logic*, 20 (1955), 151–163.
- [6] J. W. Addison and S. C. Kleene, A note on function quantification, *Proc. Amer. Math. Soc.*, 8 (1957), 1002–1006.
- [7] D. L. Kreider and H. Rogers, Jr., Constructive versions of ordinal number classes, *Trans. Amer. Math. Soc.*, 100 (1961), 325–369.
- [8] H. Putnum, Uniqueness ordinals in higher constructive number classes, *Essays on the Foundations of Mathematics*, Magnes Press, Jerusalem, 1961, and North-Holland, 1962, 190–206.
- [9] W. Richter, Extensions of the constructive ordinals, *J. Symbolic Logic*, 30 (1965), 193–211.
- [10] W. Richter, Constructively accessible ordinal numbers, *J. Symbolic Logic*, 33 (1968), 43–55.

82 (XIII.18) Contact Transformations

A. General Remarks

A transformation of $2n + 1$ variables z, x_j, p_j ($j = 1, 2, \dots, n$),

$$\begin{aligned}
 Z &= Z(z, x_1, \dots, x_n, p_1, p_2, \dots, p_n), \\
 X_j &= X_j(z, x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n), \\
 &\quad j = 1, 2, \dots, n, \\
 P_j &= P_j(z, x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n), \\
 &\quad j = 1, 2, \dots, n, \quad (1)
 \end{aligned}$$

is called a **contact transformation** in the $(n+1)$ -dimensional space \mathbf{R}^{n+1} with the coordinate system (z, x_1, \dots, x_n) if the \dagger total differential equation

$$dz - p_1 dx_1 - p_2 dx_2 - \dots - p_n dx_n = 0 \quad (2)$$

is invariant under the transformation, i.e., if the equality

$$\begin{aligned}
 dZ - P_1 dX_1 - P_2 dX_2 - \dots - P_n dX_n \\
 = \rho(dz - p_1 dx_1 - p_2 dx_2 - \dots - p_n dx_n) \quad (3)
 \end{aligned}$$

holds identically for a suitable nonzero function ρ of z, x_j, p_j . Here we assume that (1) has an inverse transformation. Using **Lagrange's bracket**

$$[f, g] = \sum_{j=1}^n \left(\frac{\partial f}{\partial p_j} \left(\frac{dg}{dx_j} \right) - \frac{\partial g}{\partial p_j} \left(\frac{df}{dx_j} \right) \right),$$

$$\left(\frac{df}{dx_j} \right) \equiv \frac{\partial f}{\partial x_j} + p_j \frac{\partial f}{\partial z},$$

we see that (1) is a contact transformation if and only if $[X_j, X_k] = [X_j, Z] = [P_j, P_k] = 0$, $[P_j, X_k] = \rho \delta_{jk}$, $[P_j, Z] = \rho P_j$, where δ_{jk} is \dagger Kronecker's delta.

From this fact it follows that the composite of two contact transformations and the inverse transformation of a contact transformation are also contact transformations. Since the identity transformation $Z = z$, $X_j = x_j$, $P_j = p_j$ is a contact transformation, the set of all contact transformations forms an infinite-dimensional \dagger topological group. Given a set of scalars p_1, \dots, p_n , a pair consisting of a point (z, x_j) and an n -dimensional hyperplane $z^* - z = \sum_{j=1}^n p_j(x_j^* - x_j)$ in an $(n+1)$ -dimensional space is called a **hypersurface element**, and the set of hypersurface elements satisfying (2) is called a **union of hypersurface elements**. Using these concepts, we can state that a transformation (1) of coordinates z, x_j, p_j ($j = 1, 2, \dots, n$) is a contact transformation if it transforms each union of hypersurface elements into another one. Consequently, if two n -dimensional hypersurfaces are tangent at a point (z, x_j) in the $(n+1)$ -dimensional space, their images under a contact transformation, which are again two n -dimensional hypersurfaces, are tangent at the image point of (z, x_j) . The name "contact transformation" is derived from this fact.

For instance, the \dagger correlation with respect to a hypersurface of the second order gives a

contact transformation. In fact, from the relation between \dagger poles and \dagger polar lines with respect to the parabola $x^2 + 2y = 0$ in a plane, we have **Legendre's transformation** $X = -p$, $Y = xp - y$, $P = -x$ ($\rho = -x$).

In general, an invertible transformation defined by the three relations $\Omega(x, y, X, Y) = 0$, $\partial\Omega/\partial X + P\partial\Omega/\partial Y = 0$, $\partial\Omega/\partial x + p\partial\Omega/\partial y = 0$ derived from a function $\Omega(x, y, X, Y)$ is a contact transformation. The function Ω is called the **generating function** of this transformation. In this transformation, to each point (x_0, y_0) there corresponds a curve $\Omega(x_0, y_0, X, Y) = 0$. These results are valid also in the case of several variables. For instance, in an $(n+1)$ -dimensional space, a transformation $Z = z - x_1 p_1 - \dots - x_v p_v$, $X_1 = p_1, \dots, X_v = p_v$, $X_{v+1} = x_{v+1}, \dots, X_n = x_n$; $P_1 = -x_1, \dots, P_v = -x_v$, $P_{v+1} = p_{v+1}, \dots, P_n = p_n$ represents a contact transformation. Here v is an integer between 1 and n . In the case $n=2$, $v=2$, this transformation reduces to a Legendre transformation; and in the case $n=2$, $v=1$, it is called **Ampère's transformation** (\rightarrow Appendix A, Table 15.IV).

B. Canonical Transformations

A transformation of $2n$ variables x_j, p_j ($j = 1, 2, \dots, n$)

$$\begin{aligned}
 X_j &= X_j(x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n), \\
 P_j &= P_j(x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n), \\
 &\quad j = 1, 2, \dots, n, \quad (4)
 \end{aligned}$$

is called a **canonical transformation** if the \dagger differential form $\sum_{j=1}^n (P_j dX_j - p_j dx_j)$ is \dagger exact in x_j, p_j , i.e., if there exists a function U of x, p such that

$$\sum_{j=1}^n (P_j dX_j - p_j dx_j) = dU. \quad (5)$$

Let (1) denote a contact transformation, and let λ denote a new variable which is different from zero. Set $\bar{x}_0 = z$, $\bar{p}_0 = \lambda$, $\bar{x}_j = x_j$, $\bar{p}_j = -\lambda p_j$ ($j = 1, 2, \dots, n$) and define a transformation of $(2n+2)$ variables by

$$\begin{aligned}
 \bar{X}_0(\bar{x}, \bar{p}) &= Z(z, x, p), \\
 \bar{P}_0(\bar{x}, \bar{p}) &= \bar{P}_0/\rho(z, x, p), \quad \bar{X}_j(\bar{x}, \bar{p}) = X_j(x, p), \\
 \bar{P}_j(\bar{x}, \bar{p}) &= -\bar{p}_0 P_j(z, x, p)/\rho(z, x, p), \\
 &\quad j = 1, 2, \dots, n, \quad (6)
 \end{aligned}$$

where x, p, \bar{x} , and \bar{p} denote points (x_1, x_2, \dots, x_n) , (p_1, p_2, \dots, p_n) , $(\bar{x}_0, \bar{x}_1, \dots, \bar{x}_n)$, and $(\bar{p}_0, \bar{p}_1, \dots, \bar{p}_n)$, respectively. Then the identity (3) when multiplied by λ on both sides becomes

$$\begin{aligned}
 \bar{P}_0 d\bar{X}_0 + \bar{P}_1 d\bar{X}_1 + \dots + \bar{P}_n d\bar{X}_n \\
 = \bar{p}_0 d\bar{x}_0 + \bar{p}_1 d\bar{x}_1 + \dots + \bar{p}_n d\bar{x}_n. \quad (7)
 \end{aligned}$$

Contact Transformations

Thus the transformation (6) represents a canonical transformation ($U = \text{constant}$). Moreover, from the definition the relation (6) is homogeneous, i.e., for all $\mu \neq 0$,

$$\bar{X}_j(\bar{x}, \mu\bar{p}) = \bar{X}_j(\bar{x}, \bar{p}), \quad \bar{P}_j(\bar{x}, \mu\bar{p}) = \mu\bar{P}_j(\bar{x}, \bar{p}),$$

$$j=0, 1, \dots, n.$$

Conversely, let a transformation of $(2n+2)$ variables

$$\bar{X}_j = \bar{X}_j(\bar{x}, \bar{p}), \quad \bar{P}_j = \bar{P}_j(\bar{x}, \bar{p}),$$

$$j=0, 1, \dots, n, \quad (8)$$

be a **homogeneous canonical transformation**.

Define a transformation of $(2n+1)$ variables z, x_j, p_j ($j=1, 2, \dots, n$) by

$$Z(z, x, p) = \bar{X}_0(z, x_1, \dots, x_n, \lambda, -\lambda p_1, \dots, -\lambda p_n),$$

$$1/\rho(z, x, p)$$

$$= \frac{1}{\lambda} \bar{P}_0(z, x_1, \dots, x_n, \lambda, -\lambda p_1, \dots, -\lambda p_n),$$

$$X_j(z, x, p) = \bar{X}_j(z, x_1, \dots, x_n, \lambda, -\lambda p_1, \dots, -\lambda p_n),$$

$$P_j(z, x, p)$$

$$= -\frac{\rho}{\lambda} \bar{P}_j(z, x_1, \dots, x_n, \lambda, -\lambda p_1, \dots, -\lambda p_n),$$

$$j=1, 2, \dots, n. \quad (9)$$

Since the necessary and sufficient condition for (8) to be a homogeneous canonical transformation is that (7) hold, it follows that

$$\frac{\lambda}{\rho} (dZ - P_1 dX_1 - \dots - P_n dX_n)$$

$$= \lambda(dz - p_1 dx_1 - \dots - p_n dx_n).$$

Thus the transformation (9) represents a contact transformation. Therefore the most general contact transformation of $(2n+1)$ variables and the most general homogeneous canonical transformation are identical concepts, differing only in the choice of notation.

The necessary and sufficient condition for (4) to be a canonical transformation is expressed by the relations

$$(X_j, X_k) = 0, \quad (X_j, P_k) = \delta_{jk}, \quad (P_j, P_k) = 0,$$

$$j, k = 1, 2, \dots, n, \quad (10)$$

where (\cdot, \cdot) is **Poisson's bracket**, i.e., for a pair of functions $f(x, p), g(x, p)$,

$$(f, g) = \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j} \frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial x_j} \right).$$

Denote Poisson's bracket in the coordinate system $X_1, X_2, \dots, X_n, P_1, P_2, \dots, P_n$ by $(\cdot, \cdot)'$, and for $f(x, p)$ and $g(x, p)$ denote their transformations by $f'(X, P)$ and $g'(X, P)$. Then the relation (10) is equivalent to

$$(f, g) = (f', g') \quad \text{for all } f \text{ and } g.$$

From this fact it follows that the set of all canonical transformations forms an infinite-dimensional topological group.

Suppose that we can take $x_1, x_2, \dots, x_n, X_1, X_2, \dots, X_n$ as $2n$ independent variables. Denote by $\Omega_1(x, X)$ the function U of (5) represented by x, X . By means of this Ω_1 , $2n$ functions which give a canonical transformation (4) are derived from the relations

$$p_j = -\frac{\partial \Omega_1}{\partial x_j}, \quad P_j = \frac{\partial \Omega_1}{\partial X_j}, \quad j=1, 2, \dots, n.$$

Conversely, for $\Omega(x, X)$ such that $\det \left(\frac{\partial^2 \Omega}{\partial x_j \partial X_k} \right) \neq 0$, a transformation determined by the relations

$$p_j = -\frac{\partial \Omega}{\partial x_j}, \quad P_j = \frac{\partial \Omega}{\partial X_j}, \quad j=1, 2, \dots, n,$$

represents a canonical transformation. For a general canonical transformation, $x_1, x_2, \dots, x_n, X_1, X_2, \dots, X_n$ are not necessarily $2n$ independent variables. But for a canonical transformation there exists n new variables $X_{i_1}, X_{i_2}, \dots, X_{i_k}, P_{j_1}, \dots, P_{j_{n-k}}$ such that $x_1, x_2, \dots, x_n, X_{i_1}, X_{i_2}, \dots, X_{i_k}, P_{j_1}, P_{j_2}, \dots, P_{j_{n-k}}$ are $2n$ independent variables, where $(i_1, i_2, \dots, i_k), (j_1, j_2, \dots, j_{n-k})$ is a partition of the set $(1, 2, \dots, n)$ to two disjoint subsets. Denote by Ω_2 the function $U - \sum_{i=1}^{n-k} X_{i_i} P_{j_i}$. Then the transformation (4) is derived from relations

$$p_j = -\frac{\partial \Omega_2}{\partial x_j}, \quad j=1, 2, \dots, n,$$

$$X_{j_l} = -\frac{\partial \Omega_2}{\partial P_{j_l}}, \quad l=1, 2, \dots, n-k,$$

$$P_{i_l} = \frac{\partial \Omega_2}{\partial X_{i_l}}, \quad l=1, 2, \dots, k.$$

These functions Ω_1, Ω_2 are called the **generating functions** of a canonical transformation.

Consider a transformation of $2n$ variables depending smoothly on a parameter t :

$$X_j = X_j(t, x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n),$$

$$P_j = P_j(t, x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n),$$

$$j=1, 2, \dots, n. \quad (11)$$

If (11) represents a canonical transformation of $2n$ variables for each t , then X_j and P_j , $j=1, 2, \dots, n$, satisfy Hamilton's canonical equations

$$\frac{dX_j}{dt} = \frac{\partial H}{\partial P_j}(t, X, P), \quad \frac{dP_j}{dt} = -\frac{\partial H}{\partial X_j}(t, X, P),$$

$$j=1, 2, \dots, n, \quad (12)$$

for some Hamiltonian function $H(t, x, p)$.

Conversely, for any solution X_j, P_j of Hamilton's canonical equations (12) depending on

parameters $x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_n$, if the transformation (11) is a canonical transformation for $t=0$, (11) represents a canonical transformation for each t . By a canonical transformation (4), the solutions of Hamilton's canonical equations

$$\frac{dx_j}{dt} = \frac{\partial H}{\partial p_j}(t, x, p), \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial x_j}(t, x, p), \quad j=1, 2, \dots, n, \quad (13)$$

are transformed to the solutions of

$$\frac{dX_j}{dt} = \frac{\partial K}{\partial P_j}(t, X, P), \quad \frac{dP_j}{dt} = -\frac{\partial K}{\partial X_j}(t, X, P), \quad j=1, 2, \dots, n, \quad (14)$$

where the new Hamiltonian function is given by $K(t, X, P) = H(t, x, p)$.

C. Applications to the Integration of Differential Equations

Contact transformations have applications to the integration of differential equations since they transform each union of surface elements into another one.

As an example, we shall describe an outline of their application to a partial differential equation of the first order

$$F(x, y, z, p, q) = 0;$$

$$p \equiv \partial z / \partial x, \quad q \equiv \partial z / \partial y. \quad (15)$$

For this purpose, we regard (15) as an equation defining unions of surface elements and transform it into a simpler equation by means of a contact transformation. If the transformed equation can be solved, then the solution of the original equation can be obtained by means of the inverse transformation. Now, let $z = \omega(x, y, a, b)$ be a †complete solution of (15). Then (15) is reduced to $Z - c = 0$ by the transformation generated by the function

$$\Omega \equiv Z - z + \omega(x, y, X, Y) - c = 0,$$

where c is a constant. In this equation the solution $X = a, Y = b, Z = c, \alpha P + \beta Q = 0$ (a, b, c, α, β are constants) plays an important role, and this line element will be called a **characteristic line element**. The characteristic line element satisfies equations that can be transformed by means of the inverse transformation into **Charpit's subsidiary equations** for (15):

$$\begin{aligned} \frac{dx}{\partial F / \partial p} &= \frac{dy}{\partial F / \partial q} = \frac{dz}{p \partial F / \partial p + q \partial F / \partial q} \\ &= \frac{-dp}{\partial F / \partial x + p \partial F / \partial z} = \frac{-dq}{\partial F / \partial y + q \partial F / \partial z}. \end{aligned} \quad (16)$$

Consequently, if we have $p = p(x, y, z; a), q =$

$q(x, y, z; a)$ from the solution $G(x, y, z, p, q) = a$ of (16) and $F = 0$, then the total differential equation $dz = p dx + q dy$ is †completely integrable, and the †general solution of this equation is a complete solution of (15). Also, if we know two independent solutions $G(x, y, z, p, q) = a, H(x, y, z, p, q) = b$ of (16) such that $[G, H] = 0$, we can obtain a complete integral of (15) by eliminating p, q among the three equations $F = 0, G = a, H = b$. This method is called the †Lagrange-Charpit method, which is applicable also to the equation $F(z, x_1, \dots, x_n; p_1, \dots, p_n) = 0$ (\rightarrow 322 Partial Differential Equations (Methods of Integration)).

D. Applications to Analytical Dynamics

Consider Hamilton's canonical equations (13). An example of system (13) is the equation of motion for a dynamical system which is derived from $\delta \int L(x_j, \dot{x}_j) dt = 0$. Here the x_j denote generalized coordinates, $p_j = \partial L / \partial \dot{x}_j$ and $H = \sum_{j=1}^n p_j \dot{x}_j - L$.

Let $S(x_1, x_2, \dots, x_n, X_1, X_2, \dots, X_n)$ be a solution depending on parameters X_1, X_2, \dots, X_n of the †Hamilton-Jacobi equation

$$\begin{aligned} H\left(t, x_1, x_2, \dots, x_n, -\frac{\partial S}{\partial x_1}, -\frac{\partial S}{\partial x_2}, \dots, -\frac{\partial S}{\partial x_n}\right) \\ = K(X_1, X_2, \dots, X_n) \end{aligned} \quad (17)$$

such that $\det \left(\frac{\partial^2 S}{\partial x_j \partial X_k} \right) \neq 0$. Then by a canonical transformation of a generating function S , i.e., by a transformation determined by the relations

$$p_j = -\frac{\partial S}{\partial x_j}, \quad P_j = \frac{\partial S}{\partial X_j}, \quad j=1, 2, \dots, n,$$

the system (13) is transformed to (14). Here

$$K(t, X, P) = H\left(t, x, -\frac{\partial S}{\partial x}\right) = K(X).$$

Then (14) becomes

$$\frac{dX_j}{dt} = 0, \quad \frac{dP_j}{dt} = -\frac{\partial K}{\partial X_j}(X), \quad j=1, 2, \dots, n,$$

and the solutions are

$$X_j(t) = X_j(0), \quad P_j(t) = P_j(0) - t \frac{\partial K}{\partial X_j} \Big|_{X=X(0)}$$

In particular, a transformation that makes $K \equiv 0$ is called a **transformation to an equilibrium system**. Thus solving Hamilton's canonical equations reduces to finding a †general solution of the Hamilton-Jacobi equation (17).

When the variables of the Hamilton-Jacobi equation (17) are totally separable we can obtain a general solution of (17) by †quadra-

tures. Then one can integrate by quadratures the system of equations (13). For many important problems exact solutions are obtained by this method (\rightarrow 271 Mechanics).

Furthermore, every mapping in an optical system is a canonical transformation; in optics, the quantity corresponding to S is called an **eikonal** (\rightarrow 180 Geometric Optics).

References

- [1] R. Courant and D. Hilbert, *Methods of mathematical physics II*, Interscience, 1962.
- [2] C. Carathéodory, *Calculus of variations and partial differential equations of the first order*, Holden-Day, 1965.
- [3] V. Arnold, *Les méthodes mathématiques de la mécanique classique*, Mir, 1976.
- [4] I. M. Gel'fand and S. V. Fomin, *Calculus of variations*, Prentice-Hall, 1963.

83 (V.3) Continued Fractions

A. The Notion of Continued Fractions

Let $\{b_n\}$ ($n=0, \dots, m$) and $\{c_n\}$ ($n=1, \dots, m$) be finite sequences of elements in a \dagger field F . A fraction of the form

$$b_0 + \frac{c_1}{b_1 + \frac{c_2}{b_2 + \dots + \frac{c_{m-1}}{b_{m-1} + \frac{c_m}{b_m}}}}$$

is called a **finite continued fraction**. It expresses an element in the field F unless division by 0 occurs in the process of reduction. Symbolically, it is also written in the forms

$$b_0 + \frac{c_1}{b_1} + \frac{c_2}{b_2} + \dots + \frac{c_{m-1}}{b_{m-1}} + \frac{c_m}{b_m},$$

$$b_0 + \frac{c_1}{|b_1|} + \frac{c_2}{|b_2|} + \dots + \frac{c_{m-1}}{|b_{m-1}|} + \frac{c_m}{|b_m|},$$

$$b_0 + \frac{c_1}{b_1} + \frac{c_2}{b_2} + \dots + \frac{c_{m-1}}{b_{m-1}} + \frac{c_m}{b_m},$$

$$\left[b_0, \frac{c_1}{b_1}, \frac{c_2}{b_2}, \dots, \frac{c_{m-1}}{b_{m-1}}, \frac{c_m}{b_m} \right],$$

etc., or more briefly,

$$b_0 + \left[\frac{c_n}{b_n} \right]_{n=1}^m.$$

If $\{b_n\}$ ($n=0, 1, \dots$) and $\{c_n\}$ ($n=1, 2, \dots$) are infinite sequences, the expression

$$b_0 + \frac{c_1}{b_1 + \frac{c_2}{b_2 + \dots + \frac{c_n}{b_n + \dots}}}$$

is called an **infinite continued fraction**. By analogy with the finite case, it is expressed by

$$b_0 + \frac{c_1}{b_1} + \frac{c_2}{b_2} + \dots + \frac{c_n}{b_n} + \dots$$

$$\text{or by } b_0 + \left[\frac{c_n}{b_n} \right]_{n=1}^{\infty}.$$

For an infinite continued fraction, the quantity

$$k_n = b_0 + \frac{c_1}{b_1} + \dots + \frac{c_n}{b_n}, \quad k_0 = b_0$$

is called its n th **convergent**, b_0 is called the **initial term**, and c_n and b_n ($n \geq 1$) are called the **partial numerator** and **partial denominator**, respectively. If F is a \dagger topological field (e.g., the real or complex number field) and the sequence $\{k_n\}$ of its elements converges, then the infinite continued fraction is said to **converge**, and the limit is called its **value**.

A finite or infinite continued fraction in which c_n ($n \geq 1$) are all equal to 1, b_0 is a rational integer, and b_n ($n \geq 1$) are all positive rational integers is called a **simple continued fraction**. It is expressed by $[b_0, b_1, \dots]$. b_n is often called the n th **partial quotient**. In the following paragraphs we shall mostly discuss simple continued fractions.

For a real number x we mean by $[x]$ the greatest integer not exceeding x . $[]$ is called the **Gauss symbol**. Let ω be any given real number, and put

$$\omega = b_0 + \frac{1}{\omega_1}, \quad b_0 = [\omega];$$

$$\omega_n = b_n + \frac{1}{\omega_{n+1}}, \quad b_n = [\omega_n], \quad n = 1, 2, \dots$$

Then an expansion of ω into a simple continued fraction

$$\omega = b_0 + \frac{1}{b_1} + \dots + \frac{1}{b_n} + \dots$$

is obtained. If ω is irrational, this expansion is determined uniquely. Conversely, any infinite simple continued fraction converges to an irrational number. If ω is rational, the process is interrupted at a finite step ($\omega_m = b_m$), resulting in

$$\omega = b_0 + \frac{1}{b_1} + \dots + \frac{1}{b_m}.$$

An alternative representation of a rational number by a simple continued fraction is given by replacing b_m above by $(b_m - 1) + 1/1$.

Examples of infinite simple continued fractions are

$$\frac{e^{2/p} + 1}{e^{2/p} - 1} = p + \frac{1}{3p + \dots + \frac{1}{(2n+1)p + \dots}},$$

where p is a natural number (J. H. Lambert), and

$$e = 2 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \dots + \frac{1}{1 + \frac{1}{2n + \frac{1}{1 + \dots}}}}}}$$

(L. Euler).

B. Convergents

Let the n th convergent of a simple continued fraction be expressed in the form of an irreducible fraction

$$\frac{P_n}{Q_n} = b_0 + \frac{1}{b_1 + \dots + b_n}, \quad n \geq 0,$$

and for convenience put $P_{-2} = 0$, $P_{-1} = 1$, $Q_{-2} = 1$, $Q_{-1} = 0$. Then we have the recurrence relations

$$P_n = b_n P_{n-1} + P_{n-2}, \quad Q_n = b_n Q_{n-1} + Q_{n-2}, \quad n \geq 0,$$

whence follows

$$P_n Q_{n-1} - P_{n-1} Q_n = (-1)^{n+1}, \quad n \geq -1.$$

Any simple continued fraction represents a real number ω which satisfies

$$\omega = (\omega_{n+1} P_n + P_{n-1}) / (\omega_{n+1} Q_n + Q_{n-1})$$

in terms of the notation defined in this and the previous section.

In particular, let ω be an irrational number. Then each of the fractions

$$\frac{P_n^{(k)}}{Q_n^{(k)}} = \frac{P_{n-2} + k P_{n-1}}{Q_{n-2} + k Q_{n-1}}, \quad k = 1, 2, \dots, b_n - 1,$$

which are inserted between two convergents P_{n-2}/Q_{n-2} and $P_n/Q_n = (P_{n-2} + b_n P_{n-1}) / (Q_{n-2} + b_n Q_{n-1})$, is called an **intermediate convergent**, while the original convergent P_n/Q_n is called a **principal convergent**.

If a fraction P/Q approximating an irrational number ω satisfies $|\omega - P/Q| < |\omega - p/q|$ for any other fraction p/q with $q \leq Q$, then it is said to give the **best approximation**. The fraction giving the best approximation of ω is always a principal or intermediate convergent $P_n^{(k)}/Q_n^{(k)}$ of ω with $k > b_n/2$ or $k = b_n/2$, $Q_n > Q_{n-1} \omega_n$.

The convergents satisfy the relation $P_n/Q_n - P_{n-1}/Q_{n-1} = (-1)^{n-1} Q_n Q_{n-1}$; hence the se-

quence $\{P_{2n}/Q_{2n}\}$ (resp. $\{P_{2n+1}/Q_{2n+1}\}$) is monotonically increasing (decreasing). Approximation by convergents is shown in the relations

$$\left| \omega - \frac{P_n}{Q_n} \right| \leq \frac{1}{Q_{n+1} Q_n}, \quad \left| \omega - \frac{P_n}{Q_n} \right| < \left| \omega - \frac{P_{n-1}}{Q_{n-1}} \right|,$$

$$\omega = \lim_{n \rightarrow \infty} \frac{P_n}{Q_n} = b_0 + \sum_{n=0}^{\infty} \frac{(-1)^n}{Q_{n+1} Q_n}.$$

In particular, $|\omega - P_n/Q_n| < 1/Q_n^2$. On the other hand, if P, Q are relatively prime positive integers with $|\omega - P/Q| < 1/2Q^2$, then P/Q is a convergent to ω (A. M. Legendre). There are several results concerning the measure of approximation by convergents. Any ω satisfies $|\omega - P_n/Q_n| < 1/(\sqrt{5} Q_n^2)$ for infinitely many n , while there exists an ω which satisfies $|\omega - P_n/Q_n| < 1/(\lambda Q_n^2)$ for only a finite number of n provided that $\lambda > \sqrt{5}$ (e.g., $\omega = \sqrt{5}$, A. Hurwitz); at least one of two adjacent convergents satisfies the inequality $|\omega - P_n/Q_n| < 1/(2Q_n^2)$ (K. T. Vahlen); at least one of three consecutive convergents satisfies the inequality $|\omega - P_n/Q_n| < 1/(\sqrt{5} Q_n^2)$ (E. Borel); $|\omega - P/Q| < 1/\sqrt{8} Q^2$ has infinitely many rational solutions P/Q whenever ω is not equivalent to $(\sqrt{5} + 1)/2$ (Hurwitz). One calls an irrational number ω **badly approximable** if there is a constant $c = c(\omega) > 0$ such that $|\omega - P/Q| > c/Q^2$ for every rational P/Q . ω is badly approximable precisely if the partial quotients of the continued fraction expansion of ω are bounded. In particular, real quadratic irrationals have bounded partial quotients (\rightarrow Section C; 182 Geometry of Numbers F).

C. Quadratic Irrationals

If an infinite simple continued fraction

$[b_0, b_1, \dots]$ satisfies $b_{m+k+v} = b_{m+v}$ ($v = 0, 1, 2, \dots$), it is called a **periodic continued fraction** and is denoted by the symbol

$[b_0, b_1, \dots, b_m, \dots, b_{m+k-1}]$. According as $m = 0$ or $m > 1$, the periodic continued fraction is said to be **pure** or **mixed**, and the sequence $b_m, b_{m+1}, \dots, b_{m+k-1}$ is called a **period**. In order for the continued fraction of ω to be periodic, it is necessary and sufficient that ω be a quadratic irrational number, i.e., a root of $ax^2 + bx + c = 0$ with rational integral coefficients a, b, c and nonsquare discriminant $b^2 - 4ac$ (J. L. Lagrange). In order for ω to be represented by a purely periodic continued fraction, it is necessary and sufficient that ω be an **irreducible quadratic irrational number**, i.e., that ω satisfy $\omega > 1$ and $0 > \omega' > -1$, where ω' is the conjugate root of ω (E. Galois). In order for ω to be equal to the square root of a nonsquare rational number, it is necessary and sufficient

that its continued fraction be of the form $[b_0, b_1, \dots, b_{k-1}, 2b_0]$ (Legendre).

D. Application to Pell's Equation

Let $ax - by = 1$ ($(a, b) = 1$) be a \dagger Diophantine equation of the first degree, and $a/b = [b_0, b_1, \dots, b_m] = P_m/Q_m$. Since $P_m Q_{m-1} - P_{m-1} Q_m = (-1)^{m-1}$, a solution of the equation is given by $x_0 = (-1)^{m-1} Q_m$, $y_0 = (-1)^m P_m$. The general solution is then represented in the form $x_0 + bt$, $y_0 + at$ ($t \in \mathbf{Z}$). This method of obtaining a solution is essentially the same as the method which uses the \dagger Euclidean algorithm.

Pell's equation $x^2 - Dy^2 = 1$ (D is a non-square integer > 1) was solved by Lagrange in terms of continued fractions. If the length of the period of D is k , all positive solutions of Pell's equation are given by $x = P_{2vk-1}$, $y = Q_{2vk-1}$ if k is odd, and by $x = P_{vk-1}$, $y = Q_{vk-1}$ if k is even ($v = 1, 2, \dots$), where P_n/Q_n denotes the n th convergent of the continued fraction expansion of \sqrt{D} . Incidentally, $x = P_{(2v-1)k-1}$, $y = Q_{(2v-1)k-1}$ ($v = 1, 2, \dots$) are the positive solutions of $x^2 - Dy^2 = -1$ provided that k is odd. There are no solutions of $x^2 - Dy^2 = \pm 1$ other than x_v, y_v ($v = 1, 2, \dots$) given by $(x_1 + \sqrt{D} y_1)^v = x_v + \sqrt{D} y_v$, where x_1, y_1 is the least positive solution. For instance, the least positive solution of $x^2 - 211y^2 = 1$ is $x = 278,354,373,650$, $y = 19,162,705,353$.

Lagrange made further use of continued fractions in order to obtain approximate values of roots of algebraic equations. The method is especially useful for precise computation of neighboring roots.

The theory of continued fractions may be investigated geometrically making use of lattices ([2, 3]; \rightarrow 182 Geometry of Numbers) (F. Klein, G. Humbert). For instance, a measure of approximation of P_n/Q_n to ω in Diophantine approximation is represented by the closeness of a lattice point (P_n, Q_n) to the straight line $y = \omega x$ on the plane.

E. Continued Fractions with Variable Terms

There are few results on continued fractions with variable terms. It is noteworthy that from the expansion of $\tan z$ into a continued fraction

$$\tan z = \frac{z}{1 + \frac{-z^2}{3 + \frac{-z^2}{5 + \frac{-z^2}{7 + \dots}}}}$$

(Lambert), the irrationality of π and of $\tan z$ for rational z ($\neq 0$) can be deduced (A. Pringsheim).

Among continued fractions with variable

terms, those of the form

$$[a_0, a_n z]_1^\infty = \frac{a_0}{1} + \frac{a_1 z}{1} + \dots + \frac{a_n z}{1} + \dots$$

are called **normal continued fractions**. Let the convergent of such a continued fraction be $P_n(z)/Q_n(z)$, and for convenience put $P_{-1}(z) = 0$, $Q_{-1}(z) = 1$. Then we have the recurrence formulas

$$P_n(z) = P_{n-1}(z) + a_n z P_{n-2}(z),$$

$$Q_n(z) = Q_{n-1}(z) + a_n z Q_{n-2}(z), \quad n \geq 1.$$

There are the further relations

$$P_n(z) Q_{n-1}(z) - P_{n-1}(z) Q_n(z) = (-1)^n z^n \prod_{v=0}^n a_v,$$

$$[a_0, a_n z]_1^\infty = \sum_{n=0}^{\infty} \frac{(-1)^n z^n}{Q_{n-1}(z) Q_n(z)} \prod_{v=0}^n a_v,$$

where the latter is formal. Let the \dagger power series expansion of the n th convergent of $[a_0, a_n z]_1^\infty$ be

$$\frac{P_n(z)}{Q_n(z)} = \sum_{v=0}^{\infty} b_{nv} z^v, \quad n \geq 1.$$

Then $b_{mv} = b_{nv}$ ($0 \leq v \leq m \leq n$). If $[a_0, a_n z]_1^\infty$ has a power series expansion about the origin, then

$$[a_0, a_n z]_1^\infty = \sum_{n=0}^{\infty} b_{nn} z^n.$$

If the supremum \bar{g} of $\{|a_n|\}_1^\infty$ is finite, then $[a_0, a_n z]_1^\infty$ converges uniformly for $|z| \leq (1/4)\bar{g}$, and hence it represents an \dagger analytic function which is holomorphic in $|z| < (1/4)\bar{g}$.

References

- [1] G. H. Hardy and E. M. Wright, An introduction to the theory of numbers, Oxford Univ. Press, 1938.
- [2] H. M. Stark, An introduction to number theory, Markham, 1970.
- [3] W. M. Schmidt, Diophantine approximation, Lecture notes in math. 785, Springer, 1980.
- [4] O. Perron, Die Lehre von den Kettenbrüchen, Teubner, third edition, 1954 (Chelsea, 1950).
- For functions defined by continued fractions and applications to moment problems,
- [5] H. S. Wall, Analytic theory of continued fractions, Chelsea, 1967.
- [6] A. Ya. Khinchin (Hinčin), Continued fractions, Noordhoff, 1963. (Original in Russian, 1949.)
- [7] A. N. Khovanskii (Hovanskii), The application of continued fractions and their generalizations to problems in approximation theory, Noordhoff, 1963. (Original in Russian, 1956.)

84 (X.2) Continuous Functions

A. General Remarks

The notion of continuity is defined for a mapping or a function $f: X \rightarrow Y$ from a topological space X to a topological space Y (\rightarrow 425 Topological Spaces G). In the present article, however, we are concerned mainly with the case where both X and Y are \dagger metric spaces with the distances ρ_X and ρ_Y , respectively. The most usual case is $X = \mathbf{R}^n$ (Euclidean space), $Y = \mathbf{R}$ (real numbers).

A function $f: X \rightarrow Y$ is said to be **continuous** at a point $x_0 \in X$ if for every positive number ε , we can select a suitable positive number δ (depending on ε and also on x_0) such that $\rho_X(x, x_0) < \delta$ implies $\rho_Y(f(x), f(x_0)) < \varepsilon$. This is equivalent to the condition that $x \rightarrow x_0$ implies $f(x) \rightarrow f(x_0)$ (\rightarrow 87 Convergence). We call f **continuous** (on X) if it is continuous at every point x_0 of X . If for every positive number ε , we can select a suitable positive number δ independent of x and y such that $\rho_X(x, y) < \delta$ implies $\rho_Y(f(x), f(y)) < \varepsilon$ for all $x, y \in X$, we call f **uniformly continuous** on X .

The \dagger supremum $\omega(\delta)$ of $\rho_Y(f(x), f(y))$ for $x, y \in X$ satisfying $\rho_X(x, y) < \delta$ is called the **modulus of continuity** of the function f in X . Uniform continuity means that $\omega(\delta) \rightarrow 0$ for $\delta \rightarrow 0$.

If $\omega(\delta) \leq M\delta^\alpha$ for suitable constants $M, \alpha > 0$, that is, if the inequality $\rho_Y(f(x), f(y)) \leq M(\rho_X(x, y))^\alpha$ holds for $x, y \in X$, then f is said to satisfy the **Hölder condition of order α** , also known as the **Lipschitz condition of order α** . If $\alpha = 1$, this condition is called simply the **Lipschitz condition**. A function satisfying one of these conditions is uniformly continuous. The family of functions satisfying the Lipschitz condition of order α is sometimes denoted by **Lip α** .

In general, the composite function $g \circ f: X \rightarrow Z$ is continuous if both functions $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ are continuous. If the ranges of f, g are both the real field \mathbf{R} (or the complex field \mathbf{C} , or more generally a \dagger topological field), then $f \pm g, fg$ are continuous if f and g are continuous; and f/g is continuous provided that $g(x) \neq 0$. If \mathbf{R} is the range of both f and g , then $\min(f, g)$ and $\max(f, g)$ are continuous when f and g are continuous. If X is \dagger connected (for example, an interval I in \mathbf{R}) and if f is continuous, the image $f(X)$ is also connected.

B. Continuity from One Side

In this section, we always assume that the domain X is an interval I in \mathbf{R} and f is a func-

tion from I to a metric space Y . A point x_0 of X is called a **discontinuity (point) of the first kind** of f if both limits $\lim_{x \uparrow x_0} f(x)$ and $\lim_{x \downarrow x_0} f(x)$ exist in Y and are different. Then we say also that f has a **jump** (or **gap**) at x_0 . If these two limits exist and have the same value, then f is continuous at x_0 .

We say that f has a **discontinuity of at most the first kind** at x_0 if f is continuous at x_0 or if x_0 is a discontinuity of the first kind for f . (Sometimes the phrase “discontinuity of the first kind” is used to mean “discontinuity of at most the first kind.”) A **discontinuity point** of f (i.e., a point at which f is not continuous) that is not of the first kind is called a **discontinuity point of the second kind**. When $\lim_{x \downarrow x_0} f(x) = f(x_0)$, we call f **right continuous** (or **continuous from the right**) at x_0 . In this case, $\lim_{x \uparrow x_0} f(x)$ need not exist. Replacing $x \downarrow x_0$ by $x \uparrow x_0$, we can similarly define the concept of being **left continuous** (or **continuous from the left**). If a function f has a finite number of discontinuity points of the first kind in the interval $[a, b]$ and is continuous at all other points, we call f a **piecewise continuous function** in $[a, b]$.

C. Semicontinuous Functions

In this section, we assume that the domain of the functions is a subset E of a metric space X , and that the range is the set of real numbers extended to include $\pm\infty$. Let x be a point in the closure of E . We denote by $M(x, \delta)$ and $m(x, \delta)$, respectively, the supremum and the infimum of the values of a given function f in the δ -neighborhood of x . We put

$$M(x) = \lim_{\delta \rightarrow 0} M(x, \delta), \quad m(x) = \lim_{\delta \rightarrow 0} m(x, \delta),$$

and call them the **upper limit function** and **lower limit function**, respectively. We have $-\infty \leq m(x) \leq M(x) \leq +\infty$. If $M(x_0) = f(x_0)$ at $x_0 \in E$, then f is called **upper semicontinuous** at x_0 . If $m(x_0) = f(x_0)$ at $x_0 \in E$ (i.e., if $-f$ is upper semicontinuous at x_0), then f is called **lower semicontinuous** at x_0 . The function with one of these two properties is said to be **semicontinuous** at x_0 .

Either of the following two conditions is necessary and sufficient for the function f to be upper semicontinuous at $x_0 \in E$: (1) $f(x_0) = +\infty$, or for every constant λ such that $f(x_0) < \lambda$, there exists a δ -neighborhood such that $M(x_0, \delta) < \lambda$. (2) For every sequence x_n of E converging to x_0 we have $\limsup_{n \rightarrow \infty} f(x_n) = f(x_0)$.

A function f is called **upper (lower) semicontinuous in E** if it is upper (lower) semicontinuous at every point $x \in E$. A necessary and sufficient condition for the upper semicontinu-

ity of the function $f(x)$ in E is that $\{x|f(x)<\alpha\}$ be a \dagger relatively open set in E for every real number α . We can define semicontinuity for functions on a topological space by using this latter property.

A real-valued function $f(x)$ is continuous at $x_0 \in E$ if and only if it is upper and lower semicontinuous at x_0 and $f(x_0)$ is finite. A function $f(x)$ is continuous on E if and only if it takes finite real values on E and both $\{x|f(x)<\alpha\}$ and $\{x|f(x)>\alpha\}$ are relatively open in E for any real number α . When E is \dagger compact, an upper (lower) semicontinuous function on E attains its (supremum) infimum at a point in E . In particular, a continuous function on a compact set E is bounded and assumes its maximum and minimum on E (**Weierstrass's theorem**). Furthermore, if E is connected (e.g., the interval I in \mathbf{R}), it follows from the connectedness of the image $f(E)$ that if $\alpha, \beta \in f(E)$ and γ lies between α and β , then $\gamma \in f(E)$ (**intermediate-value theorem**).

A real-valued function $f(x)$ on a set E of \mathbf{R} satisfies the Lipschitz condition if it is \dagger differentiable and the derivative is bounded. Such a function is also \dagger absolutely continuous, continuous, and of \dagger bounded variation. (For the polynomial approximation of real continuous functions \rightarrow 336 Polynomial Approximation.)

The limit function $f(x)$ of a monotone decreasing sequence of upper semicontinuous functions $f_n(x)$ is also upper semicontinuous. The limit function $f(x)$ of a uniformly converging sequence of continuous functions is continuous. (Regarding the \dagger equicontinuous family of functions \rightarrow 435 Uniform Convergence.)

D. Baire Functions

The limit function of a pointwise converging sequence of continuous functions defined on a metric space X is not necessarily continuous. R. Baire (*Ann. Mat. Pura Appl.*, (1899)) introduced the notion of Baire functions as follows: He named continuous functions the **functions of class 0**. Then he called a function that is a pointwise limit of a sequence of continuous functions a **function of at most class 1**. A function is said to be of class 1 if it is of at most class 1 and is not of class 0. He similarly defined the notion of **class n** for arbitrary natural number n .

Further, a function is said to be of at most class ω if it is a pointwise limit of a sequence of functions of class n , for a sequence of natural numbers n . A function is said to be of **class ω** if it is of at most class ω and is not of class n for any finite number n . In general, using \dagger transfinite induction, we can define the notion

of functions of **class ξ** for an arbitrary \dagger ordinal number ξ [4].

All these functions are called **Baire functions**. Actually, there is no function of class ξ for an uncountable ordinal number ξ . If X is a \dagger perfect set in Euclidean space, then there is actually a function defined on X of class ξ for an arbitrary countable ordinal number ξ . Hereafter, we shall be concerned with this case only.

If X has the cardinality of the \dagger continuum, then the set of all Baire functions on X has the cardinality of the continuum. On the other hand, the cardinality of all functions on X is actually greater than that of the continuum. Hence there exist functions that are not Baire functions on X . A function is a Baire function if and only if it is \dagger Borel measurable (H. Lebesgue). Therefore a necessary and sufficient condition for a function f to be a Baire function is that the set $\{x|f(x)>\alpha, x \in X\}$ be a Borel set for any real number α (\rightarrow 270 Measure Theory J). The limit of a countable sequence of Baire functions is also a Baire function. If $f(x)$ and $g(x)$ are of at most class α on X , then the following functions are also of at most class α : $|f(x)|$, $f(x) \pm g(x)$, $f(x) \cdot g(x)$, and $f(x)/g(x)$ (provided that $g(x) \neq 0$ on X).

The condition that a function f is of at most class 1 on X is equivalent to either of the following two conditions: (1) For any closed subset F of X , the restriction f^* of f to F has a continuity point in F . (2) For every real number α , the set $\{x|f(x)<\alpha, x \in X\}$ is an $\dagger F_\sigma$ set (Baire). In a \dagger complete metric space, a necessary and sufficient condition for a function f to be of at most class 1 is that the set of continuity points be dense in X .

For example, the **Dirichlet function**, which takes the value 1 at rational points and 0 at irrational points, is expressed as

$$\lim_{v \rightarrow \infty} \left(\lim_{k \rightarrow \infty} (\cos v! \pi x)^{2k} \right),$$

which is of class 2. A function $f(x, y)$ of two real variables that is continuous in each variable x and y separately is a function of at most class 1.

References

- [1] S. Saks, *Theory of the integral*, Warsaw, 1937.
 - [2] N. Bourbaki, *Eléments de mathématique*, Topologie générale, ch. 10, Actualités Sci. Ind., 1084b, Hermann, second edition, 1967.
- For Baire functions,
- [3] R. Baire, *Leçons sur les fonctions discontinues*, Gauthier-Villars, 1905.
 - [4] C. J. de La Vallée-Poussin, *Intégrales de Lebesgue, fonctions d'ensemble, classes de Baire*, Gauthier-Villars, second edition, 1934.

[5] W. Rudin, Principles of mathematical analysis, McGraw-Hill, second edition, 1964.

85 (VI.19) Continuous Geometry

A. General Remarks

The structure of a \dagger projective geometry is determined by the \dagger lattice (lattice-ordered set) of subspaces of the projective space. For this reason, this lattice itself is sometimes called a projective geometry. The concept of continuous geometry was introduced by J. von Neumann as an abstraction of lattice-theoretic properties from a special class of \dagger lattices (lattice-ordered sets) which he encountered in his research on \dagger operator rings in Hilbert spaces [1]. Continuous geometry contains projective geometry as a special case when the dimension is discrete; but more usually the lattices are of continuous dimension.

A **continuous geometry** is a \dagger complete and \dagger complemented \dagger modular lattice L (\rightarrow 243 Lattices F) that satisfies the following property and its dual (both called **properties of continuity**): For any element a of L and any subset W of L which is \dagger well ordered with respect to the ordering in L , we have $a \cap \sup w = \sup(a \cap w)$ ($w \in W$). The \dagger center Z of the lattice L is called the **center** of the continuous geometry L , which is said to be **irreducible** when Z has no elements other than the \dagger least element 0 and the \dagger greatest element I ; otherwise L is said to be **reducible**. A reducible continuous geometry is isomorphic to a sublattice of a \dagger direct product of irreducible continuous geometries.

On any continuous geometry L , there can be defined a function $d(x)$ whose values belong to a complete lattice-ordered linear space M and which satisfies the following four conditions: (1) $d(x) \geq 0$; (2) $d(x) = d(y)$ implies the existence of a common complement of x and y ; (3) $d(x \cup y) + d(x \cap y) = d(x) + d(y)$; (4) $\sup d(w) = d(\sup w)$ ($w \in W$) for any subset W of L which is well ordered with respect to the ordering in L . Such a function $d(x)$ is called a **dimension function** on L . Irreducibility of L is equivalent to the property that a real-valued dimension function can be introduced; in this case, if $d(w)$ takes only a finite number of values, then L is a finite-dimensional projective geometry; on the other hand, if $d(w)$ takes every number in the closed interval $[0, 1]$, then L is called a continuous geometry in the strict sense. An example of the latter can be constructed as a limit of a sequence of projective geometries of

increasing dimensions; another example is the lattice of projection operators of the \dagger von Neumann algebra of type II_1 . If a group G of \dagger automorphisms of L is given, there can be introduced a generalized dimension function which is invariant under G satisfying slightly weaker conditions (T. Iwamura [3]).

B. Representation of Continuous Geometry

A ring R is called a **regular ring** if it has a unity element and if, for any element a of R , there exists an element x in R such that $axa = a$. A continuous geometry L is isomorphic to the lattice (with \sup as its ordering) of \dagger principal left ideals of a regular ring R provided that $d(I) = n \cdot d(x)$ ($n \geq 4$) for some natural number n and some element x of L . The decomposition of R into a \dagger direct sum of ideals corresponds to the decomposition of L into a direct product of lattices. The condition that L is irreducible and finite-dimensional is equivalent to the condition that R is a matrix ring over a \dagger skew field; when these conditions hold and L is considered as a projective geometry, then the coordinates are given by this skew field. In continuous geometries, join and meet are often denoted by the symbols for sum and product, respectively, and sometimes a direct product of continuous geometries is called their direct sum. Sometimes, the requirement that a continuous geometry be complete is weakened to the requirement that it be \dagger conditionally σ -complete.

References

- [1] J. von Neumann, Continuous geometry, Princeton Univ. Press, 1960 (Lecture notes, 1935–1937).
- [2] F. Maeda, Continuous geometry (in Japanese), Iwanami, 1952; German translation, Kontinuierliche Geometrien, Springer, 1958.
- [3] T. Iwamura, On continuous geometries I, II, Japan. J. Math., 19 (1944), 57–71; J. Math. Soc. Japan, 2 (1950), 148–164.
- [4] L. A. Skornyakov, Complemented modular lattices and regular rings, Oliver & Boyd, 1964. (Original in Russian, 1961.)

86 (XIX.11) Control Theory

A. General Remarks

The classical theory of automatic control mostly deals with linear feedback control

systems with single input and single output. Mathematical structures of such systems must be, in principle, described in terms of ordinary linear differential equations with constant coefficients. Hence control engineers use block diagrams to describe systems, and operational calculus based on Laplace transforms to obtain response characteristics (\rightarrow 240 Laplace Transform). Thus the input/output relation of a system is described in terms of transfer functions. The main objectives of control theory are to ensure system stability, to maximize closed-loop response characteristics by choosing the best feedback, and to obtain a desirable transient response to an impulse or step input. One of the remarkable contributions to classical control theory is Nyquist's criterion for stability testing of linear feedback systems [1]. The test consists of plotting the Nyquist diagram of a transfer function in the frequency domain (complex plane), and differs essentially from the Routh-Hurwitz stability test for linear differential equations with constant coefficient. Classical control theory was almost complete by the end of World War II.

Revolutionary technological innovations in electronics and computers and the invention of new control instruments and systems developed after World War II have opened the way to modern control theory. Around 1960, three remarkable contributions were made concurrently; they are dynamic programming (proposed by R. E. Bellman [2]), Pontryagin's maximum principle (L. S. Pontryagin et al. [3]), and linear system theory (R. E. Kalman [4]). The first two give rise to mathematical tools to solve optimal control problems and to design optimal controllers or regulators. In contrast to the classical theory of control, optimal control problems are formulated in terms of the system of linear or nonlinear multivariable differential equations with multi-input forcing terms called control variables. This leads to the **state-space approach**, which has become ubiquitous in modern control theory. Linear system theory derives from the concepts of controllability and observability. The former was introduced by Pontryagin et al. [3] and later modified by Kalman [5], and the latter was introduced by Kalman [5]. Controllability and observability are concerned with the interrelation between internal states of a system and its inputs and outputs. Hence linear system theory, profoundly related to the state-space approach, remains a principal theme of modern control theory. More advanced control theories, such as adaptive control or learning control, are mostly described and formulated in the framework of linear system theory.

Modern control theory has stimulated

the development of cybernetics, the science of control and communication, and control theory is important in the information sciences. In fact, current control theory has many features that might be considered to belong as well to other areas, such as mathematical programming (\rightarrow 264 Mathematical Programming), operations research (\rightarrow 307 Operations Research), game theory (\rightarrow 173 Game Theory), prediction and filtering theory, digital signal processing, circuit theory, and computer or microprocessor technology.

Current control theory embodied in modern instrumentation and computer technology has a vast range of applications. We do not consider these applications here; we confine ourselves to some important theories that continue to be major topics in control theory.

B. Linear Dynamical System

Let the state of a physical process to be controlled be represented by a real n -dimensional column vector $x(t) = (x_1(t), \dots, x_n(t))^T \in \mathbb{R}^n$, where T denotes the transpose of a vector or matrix. If the state is determined by a differential system with input $u(t) = (u_1(t), \dots, u_r(t))^T \in \mathbb{R}^r$ and output $y(t) = (y_1(t), \dots, y_m(t))^T \in \mathbb{R}^m$,

$$\begin{aligned} \frac{d}{dt}x(t) &= A(t)x(t) + B(t)u(t), \\ y(t) &= C(t)x(t), \end{aligned} \quad (1)$$

then the process is called a **linear dynamical system** or simply a linear system. The dimensions of the coefficient matrices are given by

$$A(t) \in \mathbb{R}^{n \times n}, \quad B(t) \in \mathbb{R}^{n \times r}, \quad C(t) \in \mathbb{R}^{m \times n}.$$

If $A(t)$, $B(t)$, and $C(t)$ are constant matrices, the system described by

$$\begin{aligned} \dot{x} &= \frac{d}{dt}x = Ax + Bu, \\ y &= Cx \end{aligned} \quad (2)$$

is called a **linear time-invariant system**. Correspondingly, the system described by equation (1) is called a **linear time-varying system**.

Given an initial state $x(t_0)$ and input function $u(t)$ for $t \geq t_0$, the state of system (1) is represented by

$$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^t \Phi(t, \tau)B(\tau)u(\tau)d\tau,$$

where $\Phi(t, \tau)$ is a fundamental solution to the matrix differential equation

$$\frac{\partial}{\partial t}\Phi(t, \tau) = A(t)\Phi(t, \tau),$$

$$\Phi(\tau, \tau) = I \quad (n \times n \text{ identity matrix}).$$

Hence the output of the system is described by

$$y(t) = C(t)\Phi(t, t_0)x(t_0) + \int_{t_0}^t C(t)\Phi(t, \tau)B(\tau)u(\tau)d\tau.$$

It is seen from this equation that the input-output relation of the system is governed by

$$W(t, \tau) = C(t)\Phi(t, \tau)B(\tau) \in \mathbf{R}^{m \times r},$$

which is called a **weighting matrix**. As a special case, the weighting matrix for the time-invariant system (2) becomes

$$W(t, \tau) = W(t - \tau) = Ce^{A(t-\tau)}B,$$

where e^{At} is a matrix exponential function defined by

$$e^{At} = I + tA + \frac{t^2}{2!}A^2 + \frac{t^3}{3!}A^3 + \dots$$

The Laplace transform of the weighting matrix is

$$G(s) = \int_0^\infty e^{-st}W(t)dt = C(sI - A)^{-1}B,$$

which is called a **transfer function matrix** of the system (2). First we note that the weighting matrix $W(t, \tau)$ for the system (1) is invariant under any linear transformation $\tilde{x}(t) = P(t)x(t)$ of the state, and also that the transfer function matrix $G(s)$ for the system (2) is invariant under $\tilde{x} = Px$. For example, the transformation $\tilde{x} = Px$ yields

$$\dot{\tilde{x}} = PAP^{-1}\tilde{x} + PBu,$$

$$y = CP^{-1}\tilde{x}. \quad (3)$$

In fact, the transfer function matrix for the system (3) becomes equivalent to that for the system (2), that is,

$$\tilde{G}(s) = CP^{-1}(sI - PAP^{-1})^{-1}PB = C(sI - A)^{-1}B = G(s).$$

As a converse to this, there arises the problem of whether any linear time-invariant systems with the same transfer function matrix are similar to each other. To solve this problem, it is necessary to introduce the concept of controllability and observability for the system (1) or (2).

C. Controllability and Observability

The concept of controllability and observability plays a fundamental role in linear system theory. Roughly, controllability implies the possibility of steering the state from the input.

Definition of controllability: The linear system (1) is said to be controllable at time t_0 if for any state $x(t_0) \in \mathbf{R}^n$ and $x^1 \in \mathbf{R}^n$ there exists a control input $u(t)$, $t \in [t_0, t_1]$, that transfers the

state $x(t_0)$ to the state x^1 at a finite time $t_1 > t_0$. Otherwise, the system (1) is said to be uncontrollable at time t_0 .

The controllability theorem states: The system (1) is controllable at time t_0 if and only if there exists a finite $t_1 (> t_0)$ such that the matrix

$$D(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_0, t)B(t)B^T(t)\Phi^T(t_0, t)dt \quad (4)$$

is nonsingular. Similarly, the linear time-invariant system (2) is controllable if and only if the $n \times (nr)$ matrix

$$D = [B, AB, \dots, A^{n-1}B] \quad (5)$$

has rank n . The matrix D is called the controllability matrix of the system (2).

The concept of observability is dual to that of controllability. It implies the possibility of estimating the state from the output.

Definition of observability: The linear system (1) is said to be observable at t_0 if for any state $x(t_0) \in \mathbf{R}^n$ there exists a finite $t_1 (> t_0)$ such that the knowledge of the input $u(t)$ and the output $y(t)$ over the time interval $[t_0, t_1]$ suffices to determine the state $x(t_0)$. Otherwise, the system (1) is said to be unobservable at t_0 .

The observability theorem states: The system (1) is observable at time t_0 if and only if there exists a finite $t_1 (> t_0)$ such that the matrix

$$M(t_0, t_1) = \int_{t_0}^{t_1} \Phi^T(t, t_0)C^T(t)C(t)\Phi(t, t_0)dt \quad (6)$$

is nonsingular. Similarly, the linear time-invariant system (2) is observable if and only if the $n \times (nm)$ matrix

$$M = [C^T, A^T C^T, \dots, (A^T)^{n-1} C^T] \quad (7)$$

has rank n . The matrix M is called the observability matrix of the system (2).

We deal only with linear time-invariant systems for simplicity, although some of the main results to be stated here can be extended to linear time-varying systems. If the controllability matrix D has rank $n_c (< n)$, there exists a linear transformation $\bar{x} = Px$ with nonsingular matrix P which transforms the system (2) into

$$\begin{bmatrix} \dot{\bar{x}}_c \\ \dot{\bar{x}}_{c'} \end{bmatrix} = \begin{bmatrix} A_c & A_{12} \\ 0 & A_{c'} \end{bmatrix} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{c'} \end{bmatrix} + \begin{bmatrix} B_c \\ 0 \end{bmatrix} u, \quad (8)$$

$$y = [C_c \quad C_{c'}] \begin{bmatrix} \bar{x}_c \\ \bar{x}_{c'} \end{bmatrix}$$

so that the n_c -dimensional subsystem

$$\begin{aligned} \dot{\bar{x}}_c &= A_c \bar{x}_c + B_c u, \\ y &= C_c \bar{x}_c \end{aligned} \quad (9)$$

is controllable. Clearly this subsystem has the same transfer function matrix as the system (2).

If the observability matrix M has rank $n_o (< n)$, there exists a linear transformation $\bar{x} = Qx$ that transforms the system (2) into

$$\begin{bmatrix} \dot{x}_o \\ \dot{x}_{o'} \end{bmatrix} = \begin{bmatrix} A_o & 0 \\ A_{21} & A_{o'} \end{bmatrix} \begin{bmatrix} x_o \\ x_{o'} \end{bmatrix} + \begin{bmatrix} B_o \\ B_{o'} \end{bmatrix} u, \quad (10)$$

$$y = [C_o \quad 0] \begin{bmatrix} x_o \\ x_{o'} \end{bmatrix},$$

so that the n_o -dimensional subsystem

$$\begin{aligned} \dot{x}_o &= A_o x_o + B_o u, \\ y &= C_o x_o, \end{aligned} \quad (11)$$

is observable. This subsystem also has the same transfer function matrix as the system (2). Combination of these two properties yields the **canonical decomposition theorem** (found by Kalman [6]), which states: There is a nonsingular transformation $\bar{x} = Px$ by which the system (2) can be transformed into the canonical form

$$\begin{bmatrix} \dot{x}_{co} \\ \dot{x}_{co'} \\ \dot{x}_{c'o} \\ \dot{x}_{c'o'} \end{bmatrix} = \begin{bmatrix} A_{co} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix} \begin{bmatrix} x_{co} \\ x_{co'} \\ x_{c'o} \\ x_{c'o'} \end{bmatrix} + \begin{bmatrix} B_{co} \\ B_2 \\ 0 \\ 0 \end{bmatrix} u, \quad (12)$$

$$y = [C_{co} \quad 0 \quad C_3 \quad 0] \begin{bmatrix} x_{co} \\ x_{co'} \\ x_{c'o} \\ x_{c'o'} \end{bmatrix},$$

where the subvector x_{co} of \bar{x} is controllable and observable, $x_{co'}$ is controllable but not observable, $x_{c'o}$ is not controllable but observable, and $x_{c'o'}$ is neither controllable nor observable. Furthermore, the transfer function matrix of the system (2) is equal to that of the system (12), which is described by

$$G(s) = C_{co}(sI - A_{co})^{-1}B_{co},$$

that is, the transfer function matrix of the system (2) depends solely on the controllable and observable part of the state.

A linear system (2) is said to be **reducible** if and only if there exists a linear time-invariant system of smaller dimension that has the same transfer function matrix. Otherwise, the system is said to be **irreducible**. The canonical decomposition theorem implies that a linear time-invariant system is irreducible if and only if it is controllable and observable.

D. Realization Theory

Realization theory is concerned with determining (1) a linear time-invariant dynamical sys-

tem that has a prescribed **rational function matrix** or (2) a linear time-varying system that has a prescribed weighting matrix. Here we consider only the former, although there are many interesting results for the latter as well.

The rational function matrix $\tilde{G}(s)$ is said to be strictly proper if and only if $\tilde{G}(\infty) = 0$. A linear time-invariant system with a prescribed $\tilde{G}(s)$ as a transfer function matrix is called a **realization** of $\tilde{G}(s)$. A realization of $\tilde{G}(s)$ with the least possible dimension is said to be minimal. A minimal realization should be irreducible by definition. The fundamental theorem on minimal realization, which answers the question raised at the end of Section B, states: For a given strictly proper rational matrix $\tilde{G}(s)$, any two minimal realizations of $\tilde{G}(s)$ are mutually similar in the sense that there exists a linear transformation $\bar{x} = Px$ that transforms one of the systems into the other.

A **minimal realization** of a scalar transfer function

$$\tilde{g}(s) = \frac{b_1 s^{n-1} + \dots + b_{n-1} s + b_n}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n} \quad (13)$$

is given by

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & & \dots & \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u, \quad (14)$$

$$y = [b_n \quad b_{n-1} \quad b_{n-2} \quad \dots \quad b_1]x.$$

It is easy to see that this system is controllable. Hence system (14) is called a controllable canonical realization. Kalman [4] showed that this realization is observable and therefore minimal if and only if there is no common factor between the denominator and the numerator of $\tilde{g}(s)$. An observable canonical realization of $\tilde{g}(s)$ can be similarly written down as a dual of the system (14).

For a given strictly proper matrix $\tilde{G}(s)$, the degree of the least common denominator of all entries of $\tilde{G}(s)$ is called the degree of $\tilde{G}(s)$, provided that there is no common factor between the denominator and the numerator of each entry. It was shown by Kalman [7], together with a realization procedure, that the dimension of minimal realization of $\tilde{G}(s)$ is equal to the degree of $\tilde{G}(s)$. Other important realization procedures for rational function matrices were proposed by B. L. Ho and Kalman [8], W. A. Wolovich and P. L. Falb [9, 10], J. Rissanen and T. Kailath [11], and

B. Dickinson et al. [12]. Procedures for time-varying system realization for prescribed weighting matrices were presented by Kalman [4], D. C. Youla [15], and L. M. Silverman [16]. Realization problems for a given impulse response matrix were first discussed by C. A. Desoer and P. P. Varaiya [17], and a fast algorithm for discrete-time system realization was proposed by C. T. Mullis and R. A. Roberts [18].

E. State Estimation

When a linear system is observable, it is possible to construct a device, called a **state estimator**, that approximates the state vector. For the given system (2), define a linear dynamical system

$$\dot{z} = Dz + Ey + Gu,$$

$$\hat{x} = L_1 z + L_2 y \quad (15)$$

with inputs $y \in \mathbf{R}^m$ and $u \in \mathbf{R}^r$, state $z \in \mathbf{R}^q$, and output $\hat{x} \in \mathbf{R}^n$. If the system (2) is observable, $\text{rank}(C) = m$, and $q \geq n - m$, there exist matrices D, E, G, L_1, L_2 , and M that satisfy

$$MA - DM = EC,$$

$$L_1 M + L_2 C = I_n \quad (n \times n \text{ identity matrix}),$$

$$G = MB.$$

Then it follows that

$$\dot{\hat{x}}(t) - x(t) = L_1 e^{Dt} [z(0) - Mx(0)].$$

Furthermore, it is possible to choose a matrix D which is asymptotically stable, that is, every eigenvalue of D has negative real part. Thus $\hat{x}(t) \rightarrow x(t)$ as $t \rightarrow \infty$, and system (15) yields a state estimator. This result was first shown by D. G. Luenberger [19, 20], and thus the system (15) is called a **Luenberger observer**.

The **Kalman filter** proposed by Kalman [21] for discrete-time processes and by Kalman and R. S. Bucy [22] for continuous-time processes is an optimal state estimator for a stochastic system. Let

$$\dot{x}(t) = A(t)x(t) + B(t)u(t),$$

$$y(t) = C(t)x(t) + v(t), \quad (16)$$

in which $u(t)$ is a white Gaussian noise (\dagger Brownian motion) (\rightarrow 45 Brownian Motion) such that

$$Eu(t) \equiv 0, \quad Eu(t)u^T(s) = U(t)\delta(t-s),$$

and $v(t)$ is also a white Gaussian noise such that

$$Ev(t) \equiv 0, \quad Ev(t)v^T(s) = V(t)\delta(t-s),$$

where E and T denote the \dagger expectation and the transpose, respectively. In addition, it is as-

sumed that

$$Ex(t_0) = 0, \quad Ex(t_0)x^T(t_0) = X_0,$$

and that $x(t_0)$, $u(t)$, and $v(s)$ are uncorrelated. Under these conditions, the process $x(t)$ becomes a wide-sense Gauss-Markov process (J. L. Doob [23]). The **Kalman-Bucy filter** is a linear dynamical system described by

$$\dot{\hat{x}}(t) = [A(t) - L(t)C(t)]\hat{x}(t) + L(t)y(t),$$

where

$$L(t) = P(t)C^T(t)V^{-1}(t)$$

and $P(t)$ is a solution to the **matrix Riccati differential equation**

$$\begin{aligned} \dot{P}(t) = & A(t)P(t) + P(t)A^T(t) + B(t)U(t)B^T(t) \\ & - P(t)C^T(t)V^{-1}(t)C(t)P(t), \end{aligned}$$

$$P(t_0) = X_0.$$

The **Wiener filter** proposed in 1949 [24] is a special case of the Kalman filter in which the signal process is stationary and of single input and single output. Extensions to nonlinear filtering were investigated first by H. J. Kushner [25] and subsequently by W. M. Wonham [26], R. S. Liptzer and A. N. Shiryayev [27], and M. Fuzisaki et al. [28].

F. Optimal Control

Consider a nonlinear dynamical system described by a system of differential equations

$$\dot{x} = f(x, u), \quad (17)$$

where $x \in \mathbf{R}^n$ and $u \in \mathbf{R}^r$. For a given set U in \mathbf{R}^r , a piecewise continuous function $u(t)$ defined over $[t_0, t_1]$ is said to be an admissible control function if $u(t) \in U$ for every $t \in [t_0, t_1]$. When a functional

$$J = \int_{t_0}^{t_1} f_0(x(t), u(t)) dt$$

is given as a performance index of control and two points x^0 and x^1 are given in \mathbf{R}^n , the optimal control problem consists of finding an admissible function $u(t)$ that minimizes J and simultaneously transforms the state from $x(t_0) = x^0$ to $x(t_1) = x^1$. To solve this, a necessary condition for optimality was developed by L. S. Pontryagin and his colleagues [3] and termed the **maximum principle**. It states: In order for an admissible control $u(t)$ and a corresponding solution trajectory of equation (17) to be optimal, it is necessary that there exist a nonzero vector $\psi(t) = (\psi_1(t), \dots, \psi_n(t))^T$ such that (i) $\psi(t)$ satisfies

$$\dot{\psi}(t) = - \frac{\partial H(\psi(t), x(t), u(t))}{\partial x(t)},$$

where

$$H(\psi, x, u) = -f_0(x, u) + \psi^T f(x, u),$$

(ii) the function $H(\psi(t), x(t), u)$ of the variable $u \in U$ attains its maximum at the point $u = u(t)$ almost everywhere in $[t_0, t_1]$, and (iii) at the terminal time t_1 it holds that

$$H(\psi(t_1), x(t_1), u(t_1)) = 0.$$

When $f(x, u)$ is linear in x and u , $f_0(x, u) \equiv 1$, and t_1 is free, we have the **time-optimal control** problem investigated by Bellman et al. [29] and J. P. LaSalle [30]. When $f_0(x, u)$ is instead a quadratic function of x and u , we have the **optimal regulator problem** solved by Kalman [31]. The relation between the maximum principle and the calculus of variations (\rightarrow 46 Calculus of Variations) has been pointed out by L. D. Berkovitz [32]. Generalizations of the maximum principle and other results on optimal control problems are found in [33–35].

G. Miscellany

There are many other areas of control theory. For details, we refer the reader to [36] for adaptive control, [37] for stability theory, and [38] for system identification.

References

- [1] H. Nyquist, Regeneration theory, *Bell Syst. Tech. J.*, 11 (1932), 126–147.
- [2] R. E. Bellman, *Dynamic programming*, Princeton Univ. Press, 1957.
- [3] L. S. Pontryagin, V. G. Boltyanskiĭ, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, Interscience, 1962.
- [4] R. E. Kalman, Mathematical description of linear dynamical systems, *SIAM J. Control*, 1 (1963), 152–192.
- [5] R. E. Kalman, On the general theory of control systems, *Proc. First Int. Congress Automatic Control*, Butterworth, 1960, 481–493.
- [6] R. E. Kalman, Canonical structure of linear dynamical systems, *Proc. Nat. Acad. Sci. US*, 48 (1962), 596–600.
- [7] R. E. Kalman, Irreducible realization and the degree of a rational matrix, *SIAM J. Appl. Math.*, 13 (1965), 520–544.
- [8] B. L. Ho and R. E. Kalman, Effective construction of linear state variable models from input/output data, *Regelungstechnik*, 12 (1966), 545–548.
- [9] W. A. Wolovich and P. L. Falb, On the structure of multivariable systems, *SIAM J. Control*, 7 (1969), 437–451.

- [10] W. A. Wolovich, The determination of state-space representations for linear multi-variable systems, *Automatica*, 9 (1973), 97–106.
- [11] J. Rissanen and T. Kailath, Partial realization of random systems, *Automatica*, 8 (1972), 389–396.
- [12] B. Dickinson, M. Morf, and T. Kailath, A minimal realization algorithm for matrix sequences, *IEEE Trans. Automatic Control*, AC-19 (1974), 31–38.
- [13] B. D. O. Anderson, A system theory criterion for positive real matrices, *SIAM J. Control*, 5 (1967), 171–182.
- [14] B. D. O. Anderson, An algebraic solution to the spectral factorization problems, *IEEE Trans. Automatic Control*, AC-12 (1967), 410–414.
- [15] D. C. Youla, The synthesis of linear dynamical systems from prescribed weighting patterns, *SIAM J. Appl. Math.*, 14 (1966), 527–549.
- [16] L. M. Silverman, Synthesis of impulse response matrices by internally stable and passive realizations, *IEEE Trans. Circuit Theory*, CT-15 (1968), 238–245.
- [17] C. A. Desoer and P. P. Varaiya, The minimal realization of a nonanticipative impulse response matrix, *SIAM J. Appl. Math.*, 15 (1967), 754–764.
- [18] C. T. Mullis and R. A. Roberts, The use of second-order information in the approximation of discrete-time linear systems, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-24 (1976), 226–238.
- [19] D. G. Luenberger, Observing the state of a linear system, *IEEE Trans. Mil. Electron.*, MIL-8 (1964), 74–80.
- [20] D. G. Luenberger, Observers for multi-variable systems, *IEEE Trans. Automatic Control*, AC-11 (1966), 190–197.
- [21] R. E. Kalman, A new approach to linear filtering and prediction problems, *Trans. ASME, J. Basic Eng.*, 82 (1960), 35–45.
- [22] R. E. Kalman and R. S. Bucy, New results in linear filtering and prediction theory, *Trans. ASME, J. Basic Eng.*, 83 (1961), 95–108.
- [23] J. L. Doob, *Stochastic processes*, Wiley, 1953.
- [24] N. Wiener, *Extrapolation, interpolation and smoothing of stationary time series, with engineering applications*, Wiley, 1949.
- [25] H. J. Kushner, On the dynamical equations of the conditional density functions with applications to optimal stochastic control theory, *J. Math. Appl.*, 8 (1964), 332–344.
- [26] W. M. Wonham, Some applications of stochastic differential equations to optimal nonlinear filtering, *SIAM J. Control*, 2 (1965), 347–369.
- [27] R. S. Liptzer and A. N. Shiryaev, Non-

- linear filtering of Markov diffusion processes (in Russian), Trudy Mat. Inst. Steklov., 104 (1968), 135–180.
- [28] M. Fuzisaki, G. Kallianpur, and H. Kunita, Stochastic differential equations for the nonlinear filtering problem, Osaka J. Math., 9 (1972), 19–40.
- [29] R. Bellman, I. Glicksberg, and O. Gross, On the “bang-bang” control problem, Quart. Appl. Math., 14 (1956), 11–18.
- [30] J. P. LaSalle, The time optimal control problem, Contributions to the Theory of Nonlinear Oscillations V, Ann. Math. Studies, Princeton Univ. Press (1960), 1–24.
- [31] R. E. Kalman, Contributions to the theory of optimal control, Bol. Soc. Mat. Mexicana, second ser., 5 (1960), 102–119.
- [32] L. D. Berkovitz, Variational methods in problems of control and programming, J. Math. Anal. Appl., 3 (1961), 145–169.
- [33] E. B. Lee and L. Markus, Foundations of optimal control theory, Wiley, 1967.
- [34] A. G. Butkovskii, Theory of optimal control of distributed parameter systems, American Elsevier, 1969.
- [35] J. L. Lions, Optimal control of systems governed by partial differential equations, Springer, 1971.
- [36] Y. D. Landau, Adaptive control, Dekker, 1979.
- [37] N. P. Bhatia and G. P. Szegő, Stability theory of dynamical systems, Springer, 1970.
- [38] R. K. Mehra and D. G. Lainiotis (eds.), System identification: Advances and case studies, Academic Press, 1976.

87 (II.19) Convergence

A. Introduction

The notion of convergence was first introduced in the real number system to deal with sequences of numbers, functions, series, or definite integrals (\rightarrow 379 Series; 216 Integral Calculus). The notion was then extended to the case of generalized sequences where the index moves over a directed set, and the terms are in a topological space.

B. Convergence of Sequences of Numbers

A sequence $\{a_n\}$ of numbers is said to be **convergent** to a number a or to **converge** to a , written $\lim_{n \rightarrow \infty} a_n = a$ or $a_n \rightarrow a$ as $n \rightarrow \infty$, if for any positive number ε we can choose a (suffi-

ciently large) natural number n_0 such that for every n larger than n_0 the inequality $|a_n - a| < \varepsilon$ holds. Then a is called the **limit** (or **limit point**) of the sequence $\{a_n\}$. If $\{a_n\}$ has a limit, it is unique. A sequence which is not convergent is said to be **divergent** or to **diverge**.

A set A of real numbers is said to be **bounded from above** if there is a real number b such that $a \leq b$ for all $a \in A$, **bounded from below** if there is a real number c such that $a \geq c$ for all $a \in A$, and **bounded** if it is bounded from above and below. A sequence $\{a_n\}$ of real numbers is said to be **monotonically increasing** (**monotonically decreasing**), written $a_n \uparrow$ ($a_n \downarrow$), if $a_1 \leq a_2 \leq \dots \leq a_n \leq a_{n+1} \leq \dots$ ($a_1 \geq a_2 \geq \dots \geq a_n \geq a_{n+1} \geq \dots$). A monotonically increasing or decreasing sequence is called a **monotone sequence**.

C. Criteria for the Convergence of Sequences of Numbers

Every bounded monotone sequence of real numbers is convergent; its limit is $\sup\{a_n\}$ ($\inf\{a_n\}$) (\rightarrow 355 Real Numbers B) if it is monotonically increasing (decreasing). For any bounded sequence $\{a_n\}$ of real numbers, setting $\alpha_n = \inf\{a_n, a_{n+1}, \dots\}$ and $\beta_n = \sup\{a_n, a_{n+1}, \dots\}$, we have $\alpha_n \uparrow$, $\beta_n \downarrow$, and $\alpha_n \leq a_n \leq \beta_n$. Hence $\lim_{n \rightarrow \infty} \alpha_n = \alpha$ ($= \sup\{\alpha_n\}$) and $\lim_{n \rightarrow \infty} \beta_n = \beta$ ($= \inf\{\beta_n\}$) exist. α is called the **inferior limit** (or **limit inferior**) of $\{a_n\}$, written $\liminf_{n \rightarrow \infty} a_n$ or $\underline{\lim}_{n \rightarrow \infty} a_n$, while β is called the **superior limit** (or **limit superior**), written $\limsup_{n \rightarrow \infty} a_n$ or $\overline{\lim}_{n \rightarrow \infty} a_n$. If $\limsup_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} a_n = \alpha$, then $\lim_{n \rightarrow \infty} a_n$ exists and equals α . The limit of a convergent subsequence of a sequence $\{a_n\}$ of numbers is called an **accumulation point** of the sequence. Here we should distinguish an accumulation point of a sequence $\{a_n\}$ from an accumulation point of $\{a_n\}$ viewed as a set (\rightarrow 425 Topological Spaces O); for example, if $a_n = 1$ for every n , then the former is 1 and the latter does not exist. For any bounded sequence of real numbers, its superior (inferior) limit is the maximum (minimum) of its accumulation points. Moreover, if β is the superior limit of a sequence $\{a_n\}$, then for any positive number ε , there exist only a finite number of n 's for which a_n is greater than $\beta + \varepsilon$, while there may exist an infinite number of n 's for which a_n is less than $\beta - \varepsilon$. The inferior limit of the sequence has a similar property.

Suppose that we are given a sequence $\{a_n\}$ of real numbers and that there exist two sequences $\{u_n\}$ and $\{v_n\}$ such that $u_n \leq a_n \leq v_n$, $\lim(u_n - v_n) = 0$, $\{u_n\}$ is monotonically increasing, and $\{v_n\}$ is monotonically decreasing. Then $\lim a_n$ exists and is equal to $\lim u_n =$

$\lim v_n$ (**principle of nested intervals**). In particular, if $\limsup a_n = \liminf a_n$, then $\lim a_n$ exists. The converse also holds.

If $\{a_n\}$ is convergent, then $|a_n - a_m| \rightarrow 0$ as $n, m \rightarrow \infty$, and vice versa; that is, $\{a_n\}$ is convergent if and only if for any positive number ε there exists a positive integer n_0 such that $|a_n - a_m| < \varepsilon$ for all $n, m \geq n_0$ (**Cauchy's criterion**).

D. Infinity

For a set A of real numbers, the expression $\sup A = +\infty$ means that A is not bounded from above; $\inf A = -\infty$ means that A is not bounded from below. For a sequence $\{a_n\}$ of real numbers, $\lim a_n = +\infty$ means that for any real number b there exists a positive integer n_0 such that $a_n > b$ for all $n \geq n_0$; the notation $\lim a_n = -\infty$ has a similar meaning. The symbols $+\infty$ and $-\infty$ are called **positive** (or **plus**) **infinity** and **negative** (or **minus**) **infinity**, respectively. We say that the limit of $\{a_n\}$ is $+\infty$ ($-\infty$) if $\lim a_n = +\infty$ ($-\infty$). In these cases, we customarily say that $\{a_n\}$ **diverges** (or is **divergent**) to $+\infty$ ($-\infty$), or that a_n becomes positively (negatively) infinite as $n \rightarrow \infty$. We also define $\limsup a_n = +\infty$ ($\liminf a_n = -\infty$) to mean $\sup\{a_n\} = +\infty$ ($\inf\{a_n\} = -\infty$). A sequence $\{a_n\}$ is said to **oscillate** if $\limsup a_n > \liminf a_n$.

We now have the following propositions concerning sequences of numbers: If $\lim a_n = a$ and $\lim b_n = b$, then $\lim(\alpha a_n + \beta b_n) = \alpha a + \beta b$, $\lim(a_n b_n) = ab$, and $\lim(a_n/b_n) = a/b$ (provided that $b_n \neq 0, b \neq 0$). For sequences of real numbers, these formulas also hold when a or b is infinity. In those cases we set $\alpha \cdot (\pm\infty) = \pm\infty$ ($\alpha > 0$), $\alpha \cdot (\pm\infty) = \mp\infty$ ($\alpha < 0$), $\alpha \pm \infty = \pm\infty$, $\alpha/(\pm\infty) = 0$ for a real number α . The cases $0 \cdot (\pm\infty)$, $+\infty + (-\infty)$, $\pm\infty/(\pm\infty)$ are excluded.

E. Convergence of Sequences of Points in a Topological Space

A sequence $\{a_n\}$ of points in a topological space (\rightarrow 425 Topological Spaces) is said to **converge** to a point a if for any neighborhood U of a there exists a positive integer n_0 such that $a_n \in U$ for all $n \geq n_0$. The point a is called a **limit** (or **limit point**) of $\{a_n\}$ and we write $\lim_{n \rightarrow \infty} a_n = a$ or $a_n \rightarrow a$ as $n \rightarrow \infty$. A sequence is said to **diverge** if it does not converge to any point.

In particular, the set \mathbf{R} of all real numbers is a topological space in which the set of intervals $(a - \varepsilon, a + \varepsilon)$ for some $\varepsilon > 0$ is a \dagger base for the neighborhood system of a point a , so that the notion of limit in \mathbf{R} explained previously is

a special case of the same notion in a topological space. By adding the symbols $+\infty$ and $-\infty$ to \mathbf{R} , we obtain the topological space $\bar{\mathbf{R}}$, in which any set containing $\{x | x > \alpha, x \in \bar{\mathbf{R}}\}$ ($\{x | x < \alpha, x \in \bar{\mathbf{R}}\}$) for some $\alpha \in \mathbf{R}$ is a neighborhood of $+\infty$ ($-\infty$), where the ordering is defined as $-\infty < \alpha < +\infty$ ($\alpha \in \mathbf{R}$). Then $\lim a_n = +\infty$ ($-\infty$) is interpreted as convergence in the topological space $\bar{\mathbf{R}}$. The elements of $\bar{\mathbf{R}}$ are called **extended real numbers**.

In the case where the topological space is a \dagger metric space (\rightarrow 273 Metric Spaces) with metric ρ , $a_n \rightarrow a$ is equivalent to $\rho(a_n, a) \rightarrow 0$.

For convergence of sequences of points in a topological space, the following properties (S) hold: (S) (1). If $a_n = a$ for all n , then $\lim a_n = a$; (2) if $a_n \rightarrow a$, then $a_{n_k} \rightarrow a$ for any subsequence $\{a_{n_k}\}$; (3) if there is a point a such that any subsequence $\{a_{n_k}\}$ of $\{a_n\}$ has a suitable subsequence converging to a , then $a_n \rightarrow a$. In a \dagger Hausdorff space (e.g., a metric space), the additional property (S*) holds: If a sequence $\{a_n\}$ has a limit, this limit is unique, and can be denoted $\lim a_n$.

F. Limits of Functions

Let a real-valued function $f(x)$ of a real variable x be defined for $x \neq a$ belonging to a neighborhood of the point a . We say that the **limit** of $f(x)$ is b as x tends to a , and write $\lim_{x \rightarrow a} f(x) = b$ or $f(x) \rightarrow b$ as $x \rightarrow a$, if for any positive number ε there exists a positive number δ such that $0 \neq |x - a| < \delta$ implies $|f(x) - b| < \varepsilon$. Replacing $0 \neq |x - a| < \delta$ by $a < x < a + \delta$ ($a - \delta < x < a$), we define $f(x) \rightarrow b$ as $x \rightarrow a + 0$ ($x \rightarrow a - 0$) and say that b is the **limit on the right** (**left**) of $f(x)$ as x tends to a . We define $f(x) \rightarrow +\infty$ or $f(x) \rightarrow -\infty$ as $x \rightarrow a$ analogously to the case of sequences. The expression $f(x) \rightarrow b$ as $x \rightarrow +\infty$ means that for any positive number ε there is a real number k such that $|f(x) - b| < \varepsilon$ for any $x > k$. There are similar definitions for $x \rightarrow -\infty$ and $b = \pm\infty$. When $f(x) \rightarrow \pm\infty$ as $x \rightarrow a$, we often say that f **diverges** definitely at a .

In general, for a mapping f from a subset D of a topological space X into a topological space Y , with a point a in the closure of D and a point b in Y , $\lim_{x \rightarrow a} f(x) = b$ or $f(x) \rightarrow b$ as $x \rightarrow a$ means that any neighborhood V of b contains $f(U \cap D - \{a\})$ for some neighborhood U of a . If Y is a Hausdorff space, b is unique (if it exists) for given f and a . This point b is called the **limit** (or **limit value**) of $f(x)$ as $x \rightarrow a$.

It is easy to see that this definition of $\lim_{x \rightarrow a} f(x) = b$ is a generalization of the cases where the topological spaces are $\bar{\mathbf{R}}$ or \mathbf{R} . Let \mathbf{N} be the set of all natural numbers, and let $\bar{\mathbf{N}} =$

$\mathbf{N} \cup \{+\infty\}$ be supplied with the \dagger relative topology as a subspace of $\bar{\mathbf{R}}$. A sequence $\{a_n\}$ of real numbers or points can be identified with a mapping f from \mathbf{N} into \mathbf{R} ($\mathbf{N} \rightarrow \mathbf{R}$) defined by $f(n) = a_n$. It converges to a if and only if $\lim_{n \rightarrow \infty} f(n) = a$.

Suppose, in particular, that f is a mapping from a metric space (X, ρ) into a metric space (Y, σ) . Then, $f(x) \rightarrow b$ as $x \rightarrow a$ means that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $\sigma(f(x), b) < \varepsilon$ for all $x \in D$ such that $0 < \rho(x, a) < \delta$. Thus $f(x) \rightarrow b$ as $x \rightarrow a$ if and only if $f(x_n) \rightarrow b$ for any $\{x_n\}$ in D with $x_n \rightarrow a$. If we set $\rho(z_1, z_2) = |z_1 - z_2|$ for complex numbers z_1, z_2 , the function ρ supplies the set of all complex numbers \mathbf{C} with a metric and \mathbf{C} becomes isometric to the plane \mathbf{R}^2 (\rightarrow 74 Complex Numbers). Thus the cases $X = \mathbf{C}$ or $Y = \mathbf{C}$ are particular cases of the above generalization. Furthermore, we introduce the \dagger Riemann sphere $\bar{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$ by adding the \dagger point at infinity ∞ to \mathbf{C} . We can define a topology on $\bar{\mathbf{C}}$ such that any set containing $\{\infty\} \cup \{z \mid |z| > r\}$ for some positive number r is a neighborhood of ∞ . Thus we can define the notions $f(z) \rightarrow b$ as $z \rightarrow \infty$, $f(x) \rightarrow \infty$ as $x \rightarrow a$, etc., for a complex-valued function f by considering f as a mapping from the topological space $\bar{\mathbf{C}}$ into itself. Then $f(x) \rightarrow \infty$ is equivalent to $1/f(x) \rightarrow 0$.

G. Orders of Infinities and Infinitesimals

Let f be a complex-valued function defined on a topological space X and a a point of X . Then f is called an **infinity** (at a) or an **infinitesimal** (at a) if $f(x) \rightarrow \infty$ as $x \rightarrow a$ or $f(x) \rightarrow 0$ as $x \rightarrow a$, respectively. Suppose that f and g are infinities and f/g is an infinitesimal. Then f is said to be of **lower order** than g , and g is said to be of **higher order** than f . If both f/g and g/f are bounded, then f is said to be of **the same order** as g . This last relation, written $f \sim g$, is an equivalence relation. An infinity f is said to be of **the n th order** with g if $f \sim g^n$. For two infinitesimals f and g , f is called of **higher order** than g and g of **lower order** than f if f/g is an infinitesimal. For infinitesimals, the terms of **the same order** and of **the n th order** are defined similarly as above. In particular, when $X = \mathbf{C}$ and $a = \infty$, we usually omit the phrase "at ∞ ." Also, for such a function f , we customarily say that the **order of an infinity (infinitesimal)** is n if $f \sim z^n$ (z^{-n}).

To describe the order of an infinity or an infinitesimal simply, the following notions, due to E. Landau [10], are in common use. Let f and g be two functions. If $|f(x)/g(x)|$ is bounded as $x \rightarrow a$, then f is called **at most of the order of g** as $x \rightarrow a$, and we write $f(x) =$

$O(g(x))$ as $x \rightarrow a$. Second, if $f(x)/g(x)$ is an infinitesimal at a , then f is said to be of lower order than g as $x \rightarrow a$, and we write $f(x) = o(g(x))$ as $x \rightarrow a$. The symbols O, o , indicating the word "order," are called **Landau's symbols**. The notation $f(x) = h(x) + O(g(x))$ means $f(x) - h(x) = O(g(x))$. When we use the symbols O, o we should indicate clearly the phrase "as $x \rightarrow a$," which is sometimes omitted when no confusion is to be feared (e.g., for the case of a complex variable with $a = \infty$). These symbols are employed for sequences as well, to describe their behavior as $n \rightarrow \infty$.

H. Convergence of Nets (Moore-Smith Convergence) [7, 8]

Let \mathfrak{A} be a (preordered) \dagger directed set. A family of points in a set X with index set \mathfrak{A} (namely, a mapping from \mathfrak{A} to X) is called a **net** in X . A net is denoted by $\{x_\alpha\}_{\alpha \in \mathfrak{A}}$ ($\{x_\alpha\}_\mathfrak{A}$ or $\{x_\alpha\}$). A net $\{y_\beta\}_\mathfrak{B}$ in X is called a **subnet** of $\{x_\alpha\}_\mathfrak{A}$ if there exists a mapping $\varphi: \mathfrak{B} \rightarrow \mathfrak{A}$ such that (1) $y_\beta = x_{\varphi(\beta)}$ and (2) for any $\alpha_0 \in \mathfrak{A}$ there exists a $\beta_0 \in \mathfrak{B}$ such that $\beta \geq \beta_0$ implies $\varphi(\beta) \geq \alpha_0$. In particular, if \mathfrak{B} is a cofinal directed subset of \mathfrak{A} , $\{x_\beta\}_\mathfrak{B}$ is called a cofinal subnet of $\{x_\alpha\}_\mathfrak{A}$. A net is called a **universal net** if either $\{\alpha \mid x_\alpha \in Y\}$ or $\{\alpha \mid x_\alpha \in X - Y\}$ is residual in \mathfrak{A} for any subset Y of X . For any net there is a universal subnet.

For a net $\{x_\alpha\}_\mathfrak{A}$ in a topological space X , $\{x_\alpha\}$ is said to **converge** to a point x in X if for any neighborhood U of x there is an α_0 such that $\{x_\alpha \mid \alpha \geq \alpha_0\} \subset U$. Then a is called a **limit** of the net $\{x_\alpha\}$. We then write $x_\alpha \rightarrow x$ ($\alpha \in \mathfrak{A}$) (or simply $x_\alpha \rightarrow x$). The convergence of sequences of points is the special case where $\mathfrak{A} = \mathbf{N}$. The notion of convergence using nets was introduced by E. H. Moore and H. E. Smith.

Concerning this convergence, we have the following propositions (**D**): (**D**) (1) If $x_\alpha = x$ for all α , then $x_\alpha \rightarrow x$. (2) If $x_\alpha \rightarrow x$ and $\{y_\beta\}$ is a subnet of $\{x_\alpha\}$, then $y_\beta \rightarrow x$. (3) If for a net $\{x_\alpha\}$ there is a point x such that any subnet $\{y_\beta\}$ of $\{x_\alpha\}$ has a suitable subnet converging to x , then $x_\alpha \rightarrow x$. (4) Suppose that $x_\alpha \rightarrow x$ ($\alpha \in \mathfrak{A}$) and $y_{\alpha\beta} \rightarrow x_\alpha$ ($\beta \in \mathfrak{B}_\alpha$) for each α . Then let $\mathfrak{C} = \mathfrak{A} \times \prod \mathfrak{B}_\alpha$ be the direct product of directed sets with projections $p: \mathfrak{C} \rightarrow \mathfrak{A}$ and $p_\alpha: \mathfrak{C} \rightarrow \mathfrak{B}_\alpha$, and for $\gamma \in \mathfrak{C}$ define $z_\gamma = y_{\alpha\beta}$, where $\alpha = p(\gamma)$ and $\beta = p_\alpha(\gamma)$. Then $z_\gamma \rightarrow x$ ($\gamma \in \mathfrak{C}$). Furthermore, the space X is a \dagger Hausdorff space if and only if we have the condition (**D***): Any net in X has at most one limit.

A limit of $\{x_\alpha\}$ is denoted by $\lim x_\alpha$ or $\lim_{\alpha \in \mathfrak{A}} x_\alpha$. Then $x_\alpha \rightarrow x$ if and only if x is contained in the closure of any subnet $\{y_\beta \mid \beta \in \mathfrak{B}\}$ of $\{x_\alpha\}$. (We may consider this to be a definition of $x_\alpha \rightarrow x$.)

I. Convergence of Filters [9]

Let X be a set. A set Φ of subsets of X is called a **filter** if the following conditions are satisfied:

(i) $\emptyset \notin \Phi$ (\emptyset is the empty set); (ii) $A \subset B \subset X$ and $A \in \Phi$ imply $B \in \Phi$; (iii) $A, B \in \Phi$ imply $A \cap B \in \Phi$. Let \mathfrak{B} be a set of subsets of X and Φ be the collection of subsets of X such that each element A of Φ contains a subset belonging to \mathfrak{B} . If Φ is a filter, then \mathfrak{B} is called a **filter base** which **generates** Φ . \mathfrak{B} is a filter base if and only if (i) $\emptyset \notin \mathfrak{B}$; (ii) $A, B \in \mathfrak{B}$ implies that there is a $C \in \mathfrak{B}$ with $A \cap B \supset C$. A filter Φ is called an **ultrafilter** (or **maximal filter**) if there exists no filter which contains Φ properly. For any filter there exists an ultrafilter containing it. If $\{\Phi_\lambda\}_\Lambda$ is a family of filters, then the intersection $\bigcap \Phi_\lambda$ is a filter. If \mathfrak{F} is a filter base in the index set Λ , then $\bigcup_{M \in \mathfrak{F}} (\bigcap_{\lambda \in M} \Phi_\lambda)$ is also a filter.

We denote by $\mathcal{U}(x)$ the 'neighborhood system' of a point x in a topological space X . A filter Φ in X is said to **converge** to a point a , written $\Phi \rightarrow a$, if $\mathcal{U}(x) \subset \Phi$. A filter base \mathfrak{B} is said to **converge** to a if the filter generated by \mathfrak{B} converges to a .

The convergence of filters just defined has the following fundamental properties (L): (1) for a point a in X the filter $\Phi_a = \{A \mid a \in A \subset X\}$ converges to a ; (2) for two filters Φ and Ψ , $\Phi \rightarrow a$ and $\Phi \subset \Psi$ imply $\Psi \rightarrow a$; (3) if $\Phi_\lambda \rightarrow a$ for all members in a family $\{\Phi_\lambda\}$ of filters, then $\bigcap \Phi_\lambda = \Phi \rightarrow a$; (4) suppose that filters $\Phi_y \rightarrow y$ are assigned for all points y in a subset Y of X , and that we are given a filter Ψ in X which converges to a , generated by a filter base \mathfrak{B} in Y ; then $\bigcup_{B \in \mathfrak{B}} (\bigcap_{y \in B} \Phi_y) \rightarrow a$. Furthermore, the space X is Hausdorff if and only if we have the condition (L*): Each filter in X has at most one limit.

J. Relations among Various Definitions of Convergence

Convergence of sequences of points is a special case of that of nets. Properties (1), (2), and (3) of (D) imply (1), (2), and (3) of (S), respectively, and (D*) implies (S*). Consider a net $\{x_\alpha\}_{\alpha \in \mathfrak{A}}$ in X . Then the set $\{x_\alpha \mid \alpha \in \mathfrak{A}, \alpha \geq \alpha_0\} \mid \alpha_0 \in \mathfrak{A}\}$ of subsets of X is a filter base in X which generates a filter Φ , and $\Phi \rightarrow x$ if and only if $x_\alpha \rightarrow x$. In this situation, (L) implies (D), and (L*) implies (D*). Suppose that we are given a function $f: X \rightarrow Y$ with the domain D and a point $a \in X$. Let $\mathcal{U}(a)$ be the neighborhood system of a and assume that, for any $U \in \mathcal{U}(a)$, $U \cap D - \{a\} \neq \emptyset$. Then the set $\{f(U \cap D - \{a\}) \mid U \in \mathcal{U}(a)\}$ is a filter base. Let Φ be the filter generated by it. Then $f(x) \rightarrow b$ as $x \rightarrow a$ if and only if $\Phi \rightarrow b$. Consequently, the various types of convergence described previously can

be expressed by means of convergence of filters.

K. Convergence and Topology

In a topological space X , the concept of convergence of nets and that of filters can be defined. Conversely, convergence of nets in X defines a topology of X . In fact, let us assume that we are given a set X and a definition of convergence of filters which satisfies the properties (L). Then convergence of nets that satisfies (D) can be introduced as above. If \bar{A} is defined as the set of limits of all nets contained in a subset A of X , then \bar{A} satisfies the axiom of closures (\rightarrow 425 Topological Spaces), and a topology can be defined on X . Then we have the following propositions: (i) $a \in \bar{A}$ if and only if there is a net $\{x_\alpha\}$ with $x_\alpha \in A$ converging to a ; (ii) U is a neighborhood of a if and only if $x_\alpha \rightarrow a$ implies that there exists an α_0 such that $x_\alpha \in U$ for all $\alpha \geq \alpha_0$. Thus, if X is a topological space, it carries a "new" topology defined by way of convergence of nets. But this "new" topology coincides with the original one. Similarly, starting from convergence of filters (or nets), we can obtain a "new" definition of convergence of filters (or nets), which coincides with the initial one. In conclusion, defining a topology on a space X is the same thing as defining convergence of filters in X or of nets in X .

We shall describe here a few notions on topological spaces in terms of convergence. The fact that a topological space X is compact is equivalent to the fact that every universal net in X converges, and to the fact that every ultrafilter in X converges. Also equivalent is the fact that every net has a convergent subnet. A mapping f from a topological space X into a topological space Y is continuous at a if and only if one of the following conditions is satisfied: (1) for any net $\{x_\alpha\}$ in X converging to $a \in X$, we have $f(x_\alpha) \rightarrow f(a)$ in Y ; (2) for any filter Φ in X converging to $a \in X$, we have $f(\Phi) = \{f(M) \mid M \in \Phi\} \rightarrow f(a)$ in Y ; (3) $f(x) \rightarrow f(a)$ in Y as $x \rightarrow a$ in X (in the sense of the limit of a function at a).

M. Fréchet [6] gave a definition of a topology on a space using the notion of convergence as a foundation. A set is called an **L-space** (or **Fréchet L-space**) if convergence of sequences of points in it is defined so as to satisfy conditions (1) and (2) of (S) and (S*) (1906). Such convergence is called **star convergence** if it also satisfies (3) of (S), and in that case the space is called an **L*-space**. For any subset A of an L -space X , define \bar{A} as the set of all points a such that $x_n \rightarrow a$ for some sequence $\{x_n\}$ contained in A . Then the axioms

$\bar{A} \supset A$, $\overline{A \cup B} = \bar{A} \cup \bar{B}$, and $\overline{\emptyset} = \emptyset$ are satisfied, so that X is a \dagger generalized topological space (the axiom $\bar{A} = \bar{\bar{A}}$ is not necessarily satisfied). For a Hausdorff space X with the \dagger first countability axiom and convergence of sequences defined by means of its topology, the closure operation defined above gives the same topology as the initial one.

L. (o)-Convergence

A sequence $\{a_n\}$ of elements of an ordered set S is said to be (o)-**convergent** to an element a of S if there exist two sequences $\{u_n\}$ and $\{v_n\}$ such that $u_n \leq a_n \leq v_n$, $u_n \leq u_{n+1}$, and $v_n \geq v_{n+1}$, and $a = \sup u_n = \inf v_n$. When we write this $a_n \rightarrow a$, properties (1) and (2) of (S) and (S*) concerning the convergence of sequences hold. Next, a sequence $\{a_n\}$ is said to be (o)-**star convergent** to a if any subsequence of $\{a_n\}$ has a suitable subsequence which converges to a . Then (o)-star convergence satisfies the properties (S) and (S*).

For any set X the set $\mathfrak{P}(X)$ of all subsets of X is an ordered set under the inclusion relation. The fact that a sequence $\{A_n\}$ of subsets is (o)-convergent to a subset A is equivalent to:

$$A = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n.$$

The set A is also equal to $\lim A_n$, which is the \dagger limit of a sequence $\{A_n\}$ of subsets.

References

- [1] W. Rudin, Principles of mathematical analysis, second edition, McGraw-Hill, 1964.
- [2] J. L. Kelley, General topology, Van Nostrand, 1955.
- [3] G. Birkhoff, Lattice theory, Amer. Math. Soc. Colloq. Publ., revised edition, 1948.
- [4] N. Bourbaki, Eléments de mathématique III, Topologie générale, ch. 1, Actualités Sci. Ind., 1142d, Hermann, fourth edition, 1965; English translation, General topology, Addison-Wesley, 1966.
- [5] J. W. Tukey, Convergence and uniformity in topology, Ann. Math. Studies, Princeton Univ. Press, 1940.
- [6] M. Fréchet, Sur quelques points du calcul fonctionnel, Rend. Circ. Mat. Palermo, 22 (1906), 1–74.
- [7] E. H. Moore and H. L. Smith, A general theory of limits, Amer. J. Math., 44 (1922), 102–121.
- [8] G. Birkhoff, Moore-Smith convergence in general topology, Ann. Math., (2) 38 (1937), 39–56.
- [9] H. Cartan, Théorie des filtres, C. R. Acad. Sci. Paris, 205 (1937), 595–598; Filtrés et ultra-filtrés, *ibid.*, 777–779.
For Landau's symbol,
- [10] E. G. H. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, Teubner, 1909 (Chelsea, 1953).

88 (X.4) Convex Analysis

A. Convex Functions

A real-valued function $f(x)$ defined on a \dagger convex set D in a linear space over \mathbf{R} is called a **convex function** if for every $x, y \in D$ and $0 \leq \lambda \leq 1$ we have

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \quad (1)$$

The function $f(x)$ is called a **strictly convex function** if the sign \leq in (1) is replaced by $<$ for $x \neq y$ except when λ is either 0 or 1. If $-\psi(x)$ is convex (strictly convex), the function $\psi(x)$ is called a **concave function** (**strictly concave function**). The notion of convex function was introduced by J. L. W. V. Jensen [1] for the case where the domain D is an interval on the real line \mathbf{R} .

Sometimes the condition for a convex function is weakened in such a way that (1) is assumed only for $\lambda = 1/2$. However, if D is a \dagger topological linear space and f is continuous, then the weakened condition implies the original one. Hereafter, we mainly consider the case where D is an interval on the real line. In this case, a convex function $f(x)$ (in the weaker sense) is continuous in the interior of the interval if $f(x)$ is \dagger measurable or bounded from above on a set of positive measure (the latter was proved by A. Ostrowski [2]). In particular, suppose that $f(x)$ is defined in the interval I and is bounded from below. Then either $f(x)$ is continuous or its graph is dense in the set $\{(x, y) | x \in I, y \geq g(x)\}$, where $g(x)$ is a suitable convex continuous function (Hukuhara [3]). We note here that the original definition of a convex function $f(x)$ implies the continuity of $f(x)$ in the interior of the interval. In such a case, $f(x)$ always has \dagger right and \dagger left derivatives and satisfies $f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$ for $x < y$. Hence it is differentiable except for at most countably many points.

A function $f(x)$ is a continuous convex function in $a \leq x \leq b$ if and only if it is expressible in the form

$$f(x) = f(a) + \int_a^x \varphi(t) dt,$$

where $\varphi(t)$ is a monotone increasing function. If $f(x)$ is twice differentiable, then $f''(x) \geq 0$ ($a < x < b$) is a necessary and sufficient condition for $f(x)$ to be convex in (a, b) .

B. Convex Functions and Inequalities

If $f(x)$ is convex (in the original sense), we have, for $a_v > 0$,

$$f\left(\frac{\sum a_v x_v}{\sum a_v}\right) \leq \frac{\sum a_v f(x_v)}{\sum a_v}. \quad (2)$$

Similarly, we have, for $\varphi > 0$,

$$f\left(\frac{\int \varphi \psi dx}{\int \varphi dx}\right) \leq \frac{\int \varphi f(\psi) dx}{\int \varphi dx}. \quad (3)$$

The functions x^a ($a > 1$ or $a < 0$), $-x^a$ ($0 < a < 1$), $-\log x$, $x \log x$ are strictly convex for $x > 0$, and the functions x^{2n} ($n \geq 1$), $\exp x$, $\log(1 + e^x)$, $\sqrt{a^2 + x^2}$ ($a \neq 0$) are strictly convex in $-\infty < x < +\infty$. Applying the inequalities (2) or (3) to these functions, we obtain various inequalities, including the inequalities on means (\rightarrow 211 Inequalities).

A continuous convex function $f(x)$ over a topological linear space satisfying the relation $f(\alpha x) = \alpha f(x)$ for an arbitrary positive number α is called a **subadditive functional** and is often utilized in functional analysis.

C. M. Riesz's Convexity Theorem

Let $x = (\xi_1, \dots, \xi_n)$ be an n -tuple of complex numbers, and let $v \geq 0$. We put $N_v(x) = (\sum_{j=1}^n |\xi_j|^{1/v})^v$ for $v > 0$ and $N_0(x) = \sup |\xi_j|$. Let (α_{ij}) be an $m \times n$ complex matrix, $x = (\xi_1, \dots, \xi_n)$, $z = (\zeta_1, \dots, \zeta_m)$, $v \geq 0$, and $\mu \geq 0$. We put

$$M(v, \mu) = \sup_{N_v(x) \leq 1, N_\mu(z) \leq 1} \left| \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} \xi_i \zeta_j \right|.$$

Then $\log M(v, \mu)$ is a convex function of (v, μ) in the following sense: Let $0 < v_i \leq 1$, $0 < \mu_i \leq 1$, and $v_i + \mu_i \geq 1$ ($i = 1, 2$). Then $\log M((1-t)v_1 + tv_2, (1-t)\mu_1 + t\mu_2)$ is a convex function with respect to t for $0 \leq t \leq 1$ [4, 5]. These results are called **M. Riesz's convexity theorem**. Famous inequalities such as the \dagger Hölder inequality or the \dagger Minkowski inequality follow from this theorem. For example, let T be an \dagger additive operator from the \dagger function space $L_p(\Omega)$ into $L_p(\Omega)$ for all $1 \leq p \leq \infty$. If T is a continuous operator for $p = 1$ and $p = \infty$, and the norm of T is $\leq C$ for $p = 1$ and $p = \infty$, then T is continuous for all p ($1 < p < \infty$), and its norm is always $\leq C$.

D. Subdifferentials of Convex Functions

Currently, **convex analysis** is playing an important role in the study of nonlinear evolution equations. This analysis treats convex functions on infinite-dimensional spaces [9].

In convex analysis, it is sometimes convenient to consider a proper convex function φ on a Hilbert space X . A mapping φ from X into $(-\infty, +\infty]$ is called a **proper convex function** if φ is not identically equal to $+\infty$ and if

$$\varphi(\lambda x + (1-\lambda)y) \leq \lambda \varphi(x) + (1-\lambda)\varphi(y)$$

for every $x, y \in X$ and $0 \leq \lambda \leq 1$. The convex set

$$D(\varphi) = \{x \in X \mid \varphi(x) < \infty\} \subset X$$

is called the **effective domain** of φ . The **subdifferential** $\partial\varphi$ of φ is defined to be a multivalued function which assigns to each x in $D(\varphi)$ all elements $y \in X$ satisfying $\varphi(\xi) \geq \varphi(x) + \langle y, \xi - x \rangle$ for any $\xi \in X$. Then $\partial\varphi$ is monotone in X in the following sense: If $y_1 \in \partial\varphi(x_1)$ and $y_2 \in \partial\varphi(x_2)$, then

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0.$$

Under some conditions on φ , $\partial\varphi$ becomes a single-valued function, as is seen in the following example: Let A be a nonnegative self-adjoint operator in a Hilbert space X . For the \dagger fractional power \sqrt{A} of A , define φ by

$$\varphi(x) = \begin{cases} \frac{1}{2} \|\sqrt{A}x\|^2, & x \in D(\sqrt{A}), \\ +\infty, & \text{otherwise.} \end{cases}$$

Then $D(\varphi) = D(\sqrt{A})$ and $\partial\varphi = A$.

E. Convex Functions and Nonlinear Semigroups

Let φ be a lower semicontinuous proper convex function on a Hilbert space X . Since $A = -\partial\varphi$ is a maximal \dagger dissipative operator in X , it generates a nonlinear semigroup $\{T_t \mid t \geq 0\}$ on the closed convex set $\overline{D(A)}$:

$$T_t x = \lim_{\tau \downarrow 0} (I - \tau A)^{-[t/\tau]} x$$

(\rightarrow 286 Nonlinear Functional Analysis, 378 Semigroups of Operators and Evolution Equations). It can be shown that for any $a \in \overline{D(A)}$, $T_t a \in D(A)$ whenever $t > 0$. Thus the \dagger abstract Cauchy problem

$$\frac{du}{dt} + \partial\varphi(u) \ni 0 \quad (t > 0),$$

$$u(+0) = a,$$

is considered to be "parabolic." Furthermore,

we have

$$\left\| \frac{d}{dt} T_t a \right\| \leq \frac{1}{t} \|a - T_t a\| \quad (t > 0)$$

for all $a \in \overline{D(A)}$ (H. Brézis [10]).

References

- [1] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Math.*, 30 (1906), 175–193.
- [2] A. Ostrowski, Mathematische Miszellen XIV, Über die Funktionalgleichung der Exponentialfunktion und verwandte Funktionalgleichungen, *Jber. Deutsch. Math. Verein.*, 38 (1929), 54–62.
- [3] M. Hukuhata, Sur la fonction convexe, *Proc. Japan Acad.*, 30 (1954), 683–685.
- [4] M. Riesz, Sur les maxima des formes bilinéaires et sur les fonctionelles linéaires, *Acta Math.*, 49 (1927), 465–497.
- [5] G. O. Thorin, An extension of a convexity theorem due to M. Riesz, *Medd. Lunds Univ. Mat. Sem.*, 4 (1939), 1–5.
- [6] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, ch. 3, Cambridge Univ. Press, 1934, revised edition, 1952.
- [7] N. Bourbaki, *Eléments de mathématique*, *Espaces vectoriels topologiques*, ch. 2, *Actualité Sci. Ind.*, 1189a, Hermann, second edition, 1966.
- [8] F. A. Valentine, *Convex sets*, McGraw-Hill, 1964.
- [9] V. Barbu, Nonlinear semigroups and differential equations in Banach spaces, *Noordhoff*, 1976.
- [10] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, *Math. Studies* 5, North-Holland, 1973.

89 (VI.11) Convex Sets

A. General Remarks

A nonempty subset X of the n -dimensional Euclidean space \mathbf{R}^n is called a **convex set** if for any elements x, y in X and any number a such that $0 \leq a \leq 1$, the element $ax + (1-a)y$ of \mathbf{R}^n is also contained in X . The \dagger interior and the \dagger closure of a convex set are also convex. A point x of a convex set X is called an **extreme point** of X if x cannot be expressed as $(x_1 + x_2)/2$ in terms of a pair of distinct points x_1, x_2 in X . A bounded closed convex set is called a **convex body** if it has \dagger interior points.

Given an arbitrary nonempty subset X of \mathbf{R}^n , the minimum convex set containing X exists, called the **convex hull** of X and denoted by $[X]$. Each point x of $[X]$ can be expressed as $x = \sum_{i=1}^{n+1} a_i x_i$, where x_i belongs to X and the a_i are nonnegative numbers such that $\sum_{i=1}^{n+1} a_i = 1$. When X is a finite set, $[X]$ is called a **convex polyhedron**. If \dot{X} denotes the set of extreme points (also called **vertices**) of a convex polyhedron X , then $X = [\dot{X}]$.

For elements x, y of \mathbf{R}^n , denote the inner product by (x, y) . Given a nonzero element v of \mathbf{R}^n and a fixed number a , the \dagger hyperplane $H = \{x | (v, x) = a\}$ divides the space \mathbf{R}^n into two \dagger half-spaces $\{x | (v, x) \leq a\}$ and $\{x | (v, x) \geq a\}$, each of which is a closed convex set. If a convex set X is contained in one of the half-spaces S determined by the hyperplane H and the boundaries of X and H intersect, then we say that H is a **supporting hyperplane** of X and S is a **supporting half-space** of X . A closed convex set X is the intersection of its supporting half-spaces. A boundary point of a convex set X is contained in a supporting hyperplane of X . Given mutually disjoint convex sets X and Y , the **separation theorems** (1) and (2) hold.

(1) If X has inner points, then there exist a nonzero element v of \mathbf{R}^n and a number a such that X is contained in the set $\{x | (x, v) \geq a\}$ and Y is contained in the set $\{x | (x, v) \leq a\}$.

(2) If X and Y are closed and X is bounded, then we can replace the signs \leq and \geq in (1) by $<$ and $>$, respectively (when the separation of convex sets X and Y is described by strict inequalities, we say that X and Y are **strongly separated**).

As an immediate consequence of the separation theorems, we obtain the following proposition: Suppose that A is an $m \times n$ matrix with real entries. For an element z in a Euclidean space, we write $z \geq 0$ (> 0) if each component of z is ≥ 0 (> 0). Now if $\dagger Ay > 0$ never holds for an m -dimensional vector $y > 0$, then there exists a nonzero n -dimensional vector $x \geq 0$ such that $Ax \leq 0$ (\rightarrow 173 Game Theory).

The definitions given previously for subsets of \mathbf{R}^n can be naturally extended to the case of \dagger real topological linear spaces (\rightarrow Section G). Also, in the theory of analytic functions of several complex variables, various notions of convexity of the subsets of \mathbf{C}^n are considered (\rightarrow 21 Analytic Functions of Several Complex Variables).

B. Helly's Theorem

Suppose that we are given an index set Λ of cardinality greater than $n + 1$, and bounded closed convex sets C_λ ($\lambda \in \Lambda$) in \mathbf{R}^n . If any $n + 1$

sets of the C_λ have nonempty intersection, then the intersection of all the C_λ is nonempty (**Helly's theorem**).

This theorem has a wide range of applications. For example, we have propositions (1)–(4).

(1) If a convex set X of \mathbf{R}^n is covered by a finite number of half-spaces, then X can be covered by no more than $n+1$ half-spaces among them. (2) Let X and Y be finite subsets of \mathbf{R}^n . X and Y are strongly separated by a hyperplane if for an arbitrarily chosen subset S of $X \cup Y$ consisting of at most $n+2$ points, the sets $S \cap X$ and $S \cap Y$ are strongly separated by a hyperplane. (3) If the \dagger diameter of a subset X of \mathbf{R}^n is not greater than 2, then X is contained in a \dagger ball of radius $(2n/(n+1))^{1/2}$. (4) Let X be a convex body in \mathbf{R}^n . There exists a point x in X such that $\|x-u\|/\|v-u\| \leq n/(n+1)$, where u, v are points of intersection of an arbitrary straight line passing through x with the boundary of X , and $\|x\|$ denotes the length $(x, x)^{1/2}$ of x .

Helly's theorem can also be applied to problems of approximation of functions.

C. Ovals

The boundaries of convex bodies in \mathbf{R}^2 and \mathbf{R}^3 are called **ovals** and **ovaloids**, respectively. An oval E is a \dagger Jordan curve $C(t)$ which admits at every point $P_0 = C(t_0)$ left and right "tangents" $l_{p_0}^+$ and $l_{p_0}^-$, where $l_{p_0}^\pm$ is a straight line expressed as the set of points $P(\lambda)$, $\lambda \in \mathbf{R}$ such that $P(\lambda) - P_0 = \lambda a^\pm$, with $a^\pm = \lim_{t \rightarrow \pm 0} (C(t) - P_0)/|C(t) - P_0|$. There may exist exceptional points P on E for which left and right tangents do not coincide, but the set of such points is at most countable. Each tangent $l_{p_0}^\pm$ shares a point or a segment with E . A line satisfying this condition is called a **supporting line** of the oval. It is also a supporting hyperplane in \mathbf{R}^2 of the convex body $[E]$ in the sense of Section A. If we fix an interior point O of a convex body X and take an arbitrary point P different from O , then the boundary E of X admits one and only one supporting line $l(P)$ which is perpendicular to the line OP and meets the half-line OP . Take a rectangular coordinate system with the origin O , and let (x, y) denote the coordinates of P . Then the points (ξ, η) on $l(P)$ satisfy the equation $\xi x + \eta y = H(x, y)$, where $H(x, y)$ is a function determined for all (x, y) in \mathbf{R}^2 and satisfying the following conditions: (i) $H(0, 0) = 0$; (ii) $H(tx, ty) = tH(x, y)$, for $t \geq 0$; (iii) $H(x_1 + x_2, y_1 + y_2) \leq H(x_1, y_1) + H(x_2, y_2)$. The function $H(x, y)$ is called the **supporting line function** of E . The magnitude and shape of E are determined by H , and any function satisfying con-

ditions (i)–(iii) is a supporting line function of an oval. An oval E has a finite length $L = L(E)$, and the convex body $[E]$ has a finite area $F = F(E)$. If OP' denotes the half-line with direction opposite to that of OP and l' denotes the supporting line $l(P')$, the distance between the parallel lines l and l' is called the **breadth** of E in the direction PP' . Let $D = D(E)$ and $\Delta = \Delta(E)$ be the maximum and minimum of the breadth of E , respectively. D is the \dagger diameter of E (or $[E]$), and Δ is called the **thickness** of E (or $[E]$). In particular, if $D = \Delta$, then the oval E is called a **curve of constant breadth**. In the following inequalities, equality holds only when E is one of the figures mentioned in parentheses: (1) $L^2 \geq 4\pi F$ (circles, J. Steiner (1838)) (\rightarrow 228 Isoperimetric Problems); (2) $\pi D^2 \geq 4F$ (circles, L. Bieberbach (1915)); (3) $L \leq \pi D$ (curves of constant breadth, W. Blaschke (1916)); (4) $F \geq \Delta^2/\sqrt{3}$ (regular triangles, J. Pál (1921)). See T. Kubota, *Tôhoku Sci. Bull.*, I, 12, 13; *Tôhoku Math. J.*, 24, 49.

D. Linear Combinations of Ovals

Let H_1 and H_2 be supporting line functions of ovals E_1 and E_2 , and let t_1 and t_2 be positive numbers. Since the function $t_1 H_1 + t_2 H_2$ satisfies conditions (i)–(iii), given before, it is a supporting line function of an oval $E(t_1, t_2)$. In this case, we can also write $E(t_1, t_2) = t_1 E_1 + t_2 E_2$ and call it a **linear combination** of E_1 and E_2 . In particular, the oval $(E_1 + E_2)/2$ is called the **mean oval** of E_1 and E_2 . In general, there exists a quantity M , called the **mixed area** of E_1 and E_2 , such that $F(E(t_1, t_2)) = F(E_1)t_1^2 + 2Mt_1t_2 + F(E_2)t_2^2$. M does not depend on the choice of t_1 and t_2 , and $M^2 \geq F(E_1)F(E_2)$. Here, the equality holds if and only if E_1 and E_2 are homothetic and situated in a position of homothety. Furthermore, if $0 \leq t \leq 1$, then the square root of $F(E(t, 1-t))$ is a \dagger convex function of t (H. Minkowski).

E. Specific Ovals

Suppose that we are given an equilateral triangle ABC . Draw three circles C_1, C_2 , and C_3 with centers A, B , and C and radii equal to the length of the sides of ABC . The minor arcs AB, BC , and CA of the circles form an oval which is called a **Reuleaux triangle**. This oval is of constant breadth. Furthermore, given a fixed breadth D , the area $F(E)$ of an oval E of constant breadth D attains its minimum when E is a Reuleaux triangle. A Reuleaux triangle obtained from a triangle ABC revolves freely within the square of side AB and touches each side. In general, an oval which revolves touch-

ing the sides of a convex polygon from the inside is called an **inrevolvable** oval. Any such oval revolves inside some regular polygon (M. Fujiwara, S. Kakeya).

Various properties of an oval already described can be generalized to the case of a boundary of a convex body in \mathbf{R}^n . For the volumes of subsets A, B in \mathbf{R}^n and $A + B = \{x + y | x \in A, y \in B\}$ in \mathbf{R}^n , the Brunn-Minkowski inequality

$$[\text{vol}(A + B)]^{1/n} \geq [\text{vol}(A)]^{1/n} + [\text{vol}(B)]^{1/n}$$

holds.

F. Convex Cones

A nonempty subset X of \mathbf{R}^n is called a convex cone if for any elements x, y of X and a non-negative number a , ax and $x + y$ are contained in X . A convex cone is a convex set. Given any nonempty subset X of \mathbf{R}^n , the minimum convex cone $K(X)$ containing X exists. Given two convex cones X and Y , a convex cone $X + Y$, called the **sum** of X and Y , is defined as the set of elements $x + y$, where x, y are elements of X, Y . The intersection of convex cones X and Y is also a convex cone. Given a convex cone X , the subset of \mathbf{R}^n consisting of the elements y such that $(x, y) \leq 0$ for any element x in X is a convex cone which is called the **dual convex cone** (or **conjugate convex cone**) of X , denoted by X^* . If X is a finite set, $K(X)$ is called a **convex polyhedral cone**. For example, if v is a nonzero vector, then the half-line $(v) = \{x | x = av, a \geq 0\}$ or the half-space $(v)^* = \{x | (v, x) \leq 0\}$ is a convex polyhedral cone. A convex polyhedral cone is closed. A convex cone X is a convex polyhedral cone if and only if X is the sum of a finite number of half-lines. Given convex cones X and Y , we have propositions (1)–(3): (1) If $X_1 \subset X_2$, then $X_2^* \subset X_1^*$; (2) $(X_1 + X_2)^* = X_1^* \cap X_2^*$; (3) $X_1^* + X_2^* \subset (X_1 \cap X_2)^*$. If X_1 and X_2 are convex polyhedral cones, then $X_1^* + X_2^* = (X_1 \cap X_2)^*$. Generally, $X \subset (X^*)^* = X^{**}$ for a convex cone X . If X is a closed convex cone, then $X = X^{**}$. Namely, the **duality principle** holds for closed convex cones. A linear subspace of \mathbf{R}^n is a convex polyhedral cone. Also, if A is an $m \times n$ real matrix, the subsets $\{x | Ax = 0, x \geq 0\}$ and $\{x | Ax \geq 0\}$ are convex polyhedral cones. Since the duality principle holds for convex polyhedral cones, we obtain the \dagger Minkowski-Farkas theorem (i.e., if A is an $m \times n$ real matrix and v is an element of \mathbf{R}^m , then the equation $Ay = v$ has a solution $y \geq 0$ in \mathbf{R}^n if and only if $(v, x) \geq 0$ for all $x \in \mathbf{R}^m$ such that $Ax \geq 0$). (For linear inequalities \rightarrow 255 Linear Programming.)

G. Convex Sets in Function Spaces

The definitions of convex sets and convex cones in \mathbf{R}^n can be naturally extended to the case of any real linear space. Some of their properties can be generalized and applied to the case of function spaces.

(1) Let E be a \dagger locally convex real topological linear space satisfying Hausdorff's \dagger separation axiom. Let A and B be convex sets in E , and assume that B has interior points and $A \cap B$ is empty. Then A and B are separated by a hyperplane. Namely, there exists a nonzero \dagger continuous linear functional f on E such that $\sup f(A) \leq \inf f(B)$.

(2) Let E be as in (1), and let C be a convex set in E . If a boundary point x of C admits a nonzero continuous linear functional f such that $f(x) = \sup f(C)$, we call such a point x a **supporting point** of C , and f a **supporting functional** of C . If C has interior points, then any boundary point x of C is a supporting point of C .

(3) Let C be a closed convex set of a \dagger Banach space E . The set of supporting points of C is dense in its boundary.

A convex set C contained in the dual space E^* of a real topological linear space E is called a **regularly convex set** if for any f_0 in E^* not contained in C , there exists a point x_0 in E such that $\sup\{f(x_0) | f \in C\} < f_0(x_0)$.

Let E be a real topological linear space satisfying Hausdorff's separation axiom, and let C be a closed convex cone having 0 as its extreme point. Furthermore, assume that $C \cap (-C) = \{0\}$. If we set $x \leq y$ when $y - x \in C$, a partial ordering \leq is defined in E . For example, if E is \mathbf{R}^n , then the **positive orthant** $C = \{x = (x_i) | x_i \geq 0, i = 1, \dots, n\}$ satisfies these requirements, and the partial ordering $x \leq y$ defined by means of C is equivalent to the relation $x_i \leq y_i$ for all i .

Some of the properties of matrices of positive entries or \dagger integral operators whose \dagger kernel functions are positive-valued can be generalized to properties of mappings $f: E \rightarrow E$ such that $f(C) \subset C$ (\rightarrow 255 Linear Programming; for the \dagger Krein-Milman theorem \rightarrow 424 Topological Linear Spaces).

References

- [1] W. Blaschke, Kreis und Kugel, Veit, 1916 (Chelsea, 1949).
- [2] T. Bonnesen and W. Fenchel, Theorie der konvexen Körper, Erg. Math., Springer, 1934 (Chelsea, 1948).
- [3] W. Fenchel, Convex cones, sets and functions, Lecture notes, Princeton Univ. Press, 1953.

- [4] Convexity, Amer. Math. Soc. Proc. Symp. in Pure Math. VII (1963).
 [5] M. G. Kreĭn and M. A. Rutman, Linear operators leaving invariant a cone in a Banach space, Amer. Math. Soc. Transl., 26 (1950). (Original in Russian, 1948.)
 [6] H. Nikaido, Convex structures and economic theory, Academic Press, 1968.
 [7] B. Grünbaum, Convex polytopes, Interscience, 1967.
 [8] I. M. Yaglom (Jaglom) and V. G. Boltyanskii, Convex figures, Holt, Rinehart and Winston, 1961. (Original in Russian, 1951.)
 [9] G. Choquet, Lectures on analysis, vols. I, II, Benjamin, 1969.
 [10] R. T. Rockafellar, Convex analysis, Princeton Univ. Press, 1970.

90 (VI.13) Coordinates

A. General Remarks

Suppose that we are given a Euclidean plane E^2 and two lines $X'X$ and $Y'Y$ in E^2 perpendicular to each other. Let O be the point of intersection of $X'X$ and $Y'Y$. We identify each of the straight lines $X'X$ and $Y'Y$ with the set of real numbers \mathbf{R} ; the point O on each line is identified with zero. Let P be an arbitrary point in E^2 . We draw lines PQ , PR parallel to $Y'Y$, $X'X$, where Q , R are on $X'X$, $Y'Y$, respectively. Let x and y be the real numbers corresponding to Q and R . Thus we obtain a mapping sending the point P to the ordered pair (x, y) of real numbers. This mapping gives a one-to-one correspondence between the points P of E^2 and the ordered pairs (x, y) of real numbers in \mathbf{R}^2 . The numbers x and y are called the **coordinates** of P .

In general, given a set of mathematical objects, if we have a mechanism that assigns quantities to each element of the set, then such a mechanism is called a **coordinate system** on the set, and the quantities corresponding to each element are called its **coordinates**. In the previous example, the mechanism is called a rectangular coordinate system. Coordinate systems are also useful in expressing quantitative concepts by geometric ones which are intuitively easier to grasp, e.g., diagrams of train schedules and \dagger nomograms. \dagger Map projection, \dagger graphical calculation, \dagger descriptive geometry, etc., may be viewed as applications of the concept of coordinate systems.

In many cases, when we introduce a coordinate system in a space, it is determined uniquely by fixing a basic figure in the space.

In the case of a rectangular coordinate system on a plane E^2 , the basic figure consists of $X'X$ and $Y'Y$, which are called coordinate axes (the point O is called the origin). Sometimes it is convenient to consider real-valued functions f and g on \mathbf{R} and a coordinate system on the plane E^2 determined by the function that sends an arbitrary point P to $(f(x), g(y))$, where (x, y) are the coordinates of P in the rectangular coordinate system. Logarithmic papers, \dagger probability papers, and \dagger stochastic papers (binomial probability papers), etc., are constructed in this way to fit their respective purposes.

In various branches of mathematics there are many varieties of coordinate systems. In this article we deal with frames and coordinates, curvilinear coordinates, and local coordinates.

B. Frames and Coordinates

Suppose that we are given a space M and a \dagger transformation group G acting on M . It is desirable to introduce a coordinate system that best represents the geometric structure of M . Let G_* be a set of figures in M such that G acts \dagger simply transitively on G_* . Each element of G_* is called a **frame**. Utilizing each frame as basic figure, we introduce a coordinate system that is " G -invariant" in the following sense: Let $R \in G_*$, $X \in M$, and $C_R(X)$ be the coordinates of X with R as basic figure. Then the coordinate system is G -invariant if $C_R(X) = C_{gR}(gX)$ for any element g in G . If we have such a coordinate system for each $R \in G_*$, then the expressions of geometric properties of M in terms of the coordinates are independent of the choice of frames.

(1) Projective Coordinates. Let M be an n -dimensional \dagger projective space \mathbf{P}^n over a field K , and let G be the \dagger projective transformation group of \mathbf{P}^n . As a frame we can take the system of $n+1$ points (A_0, A_1, \dots, A_n) in general position. The \dagger homogeneous coordinates of an arbitrary point $X \in \mathbf{P}^n$ are given by the $(n+1)$ -tuple (x_0, x_1, \dots, x_n) satisfying the equation $X = \sum_{j=0}^n x_j A_j$, $x_j \in K$. They are called \dagger projective coordinates. In fact, if $(x_0, x_1, \dots, x_n) \neq (0, 0, \dots, 0)$, then (x_0, x_1, \dots, x_n) and $(\lambda x_0, \lambda x_1, \dots, \lambda x_n)$ ($\lambda \neq 0$) represent the same point in \mathbf{P}^n . A \dagger hyperplane π of \mathbf{P}^n is expressed as the set of points whose coordinates (x_0, x_1, \dots, x_n) satisfy a linear homogeneous equation $\sum_{j=0}^n x_j u_j = 0$, $u_j \in K$. Therefore, the hyperplane π is represented by the homogeneous coordinates (u_0, u_1, \dots, u_n) , called \dagger hyperplane coordinates of π .

(2) Affine Coordinates. Let M be an n -dimensional \dagger affine space E^n , and let G be the \dagger group of affine transformations of E^n . As a frame we can take the system $(O; \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$, where O , called the origin, is a point in E^n , and the set of vectors $\{\mathbf{e}_i\}$ is a basis of the \dagger standard vector space of E^n . Then \dagger inhomogeneous coordinates of an arbitrary point $X \in E^n$ are given by the n -tuple (x_1, x_2, \dots, x_n) , where $X = O + \sum_{i=1}^n x_i \mathbf{e}_i$. They are also called \dagger affine coordinates of X in E^n . Sometimes we replace G above by the group of \dagger equivalent affinities and consider the frames $(O; \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ such that the volume of $[\mathbf{e}_1, \dots, \mathbf{e}_n] = 1$.

Furthermore, if E^n has the structure of a \dagger Euclidean space, we sometimes replace G by the \dagger group of motions and consider a system of **rectangular coordinates** determined by an \dagger orthogonal frame, that is, a frame $(O; \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ such that the inner product $(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$, where δ is the \dagger Kronecker delta. By contrast, the general affine coordinate system of a Euclidean space is called a system of **oblique coordinates**. In this case the inner products $(\mathbf{e}_i, \mathbf{e}_j) = g_{ij}$ are invariants of Euclidean geometry, and the distance ρ between two points (x_i) and (y_i) is given by $\rho = (\sum_{i,j=1}^n g_{ij}(y_i - x_i)(y_j - x_j))^{1/2}$. We sometimes consider an oblique coordinate system satisfying $(\mathbf{e}_i, \mathbf{e}_j) = 1$ ($i = 1, \dots, n$). In such cases the angle θ_{ij} between two basis vectors \mathbf{e}_i and \mathbf{e}_j is determined by $g_{ij} = \cos \theta_{ij}$.

(3) Barycentric Coordinates. In an n -dimensional affine space E^n , we take $n+1$ linearly independent points A_0, A_1, \dots, A_n and denote the position vectors from a point O to these points by $\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_n$, respectively. Then for any point $X \in E^n$ there exists a unique set of numbers $(\lambda_0, \lambda_1, \dots, \lambda_n)$ such that $X = O + \sum_{j=0}^n \lambda_j \mathbf{a}_j$, $\sum_{j=0}^n \lambda_j = 1$. We call these numbers **barycentric coordinates** of X in E^n . They are independent of the choice of the point O .

(4) Plücker Coordinates. Let $V(n, m)$ be the set of all m -dimensional subspaces in an n -dimensional projective space \mathbf{P}^n . Then $V(n, m)$ has the structure of a \dagger Grassmann manifold. In order to introduce a coordinate system on $V(n, m)$, we fix a projective coordinate system on \mathbf{P}^n . An m -dimensional subspace $\pi \in V(n, m)$ in \mathbf{P}^n is spanned by $m+1$ independent points $B_0, B_1, \dots, B_m \in \mathbf{P}^n$. We denote projective coordinates of these points by $(b_{0j}), (b_{1j}), \dots, (b_{mj})$ and construct the determinants

$$p_{j_0 j_1 \dots j_m} = \begin{vmatrix} b_{0j_0} & b_{0j_1} & \dots & b_{0j_m} \\ \dots & \dots & \dots & \dots \\ b_{mj_0} & b_{mj_1} & \dots & b_{mj_m} \end{vmatrix}, \quad 0 \leq j_0, \dots, j_m \leq n.$$

Then the subspace π can be represented by homogeneous coordinates $(\dots, p_{j_0 j_1 \dots j_m}, \dots)$.

These coordinates are independent of the choice of $m+1$ points that span π and are called **Plücker coordinates** (or **Grassmann coordinates**) of π in $V(n, m)$. In these coordinates, the $p_{j_0 j_1 \dots j_m}$ are alternating and satisfy the **Plücker relations**

$$\sum_{k=0}^{m+1} (-1)^k p_{i_1 i_2 \dots i_{m+1} k} p_{j_0 \dots \hat{j}_k \dots j_{m+1}} = 0,$$

where \hat{j}_k means that j_k is removed. In particular, when $n=3$ and $m=1$, we have only one Plücker relation $Q: p_{01}p_{23} - p_{02}p_{13} + p_{03}p_{12} = 0$, which is a homogeneous equation of the second degree. In other words, the set $V(3, 1)$ of all lines in a 3-dimensional projective space \mathbf{P}^3 is realized as a quadric surface Q in a 5-dimensional projective space \mathbf{P}^5 that has $(p_{01}, p_{02}, p_{03}, p_{12}, p_{13}, p_{23})$ as projective coordinates. Moreover, when \mathbf{P}^3 is a complex projective space, we put

$$p_{01} = \xi_0 + i\xi_3, \quad p_{02} = \xi_1 + i\xi_4, \quad p_{03} = \xi_2 + i\xi_5, \\ p_{23} = \xi_0 - i\xi_3, \quad p_{13} = -\xi_1 + i\xi_4, \quad p_{12} = \xi_2 - i\xi_5$$

and obtain a relation

$$\xi_0^2 + \xi_1^2 + \xi_2^2 + \xi_3^2 + \xi_4^2 + \xi_5^2 = 0$$

corresponding to the Plücker relation. Thus every line in \mathbf{P}^3 can be represented by homogeneous coordinates $(\xi_0, \xi_1, \dots, \xi_5)$, which we call **Klein's line coordinates**.

(5) $(n+2)$ -Hyperspherical Coordinates. Let $(x_0, x_1, \dots, x_n, x_\infty)$ be projective coordinates in an $(n+1)$ -dimensional real projective space \mathbf{P}^{n+1} . An n -dimensional \dagger conformal space S^n is realized as a quadric hypersurface S^n in \mathbf{P}^{n+1} : $\sum_{i,j=1}^n g_{ij}x_i x_j - 2x_0 x_\infty = 0$, where (g_{ij}) is a positive definite symmetric matrix. A general point in \mathbf{P}^{n+1} represents a \dagger hypersphere of S^n . That is, a hypersphere represented by a point $X \in \mathbf{P}^{n+1}$ is realized as the intersection of S^n with the \dagger polar hyperplane of X with respect to S^n ; according as X lies outside of S^n , on S^n , or inside of S^n , it represents a real hypersphere, a point hypersphere, or an imaginary hypersphere. Therefore any hypersphere of S^n in \mathbf{P}^{n+1} is expressed by homogeneous coordinates $(x_0, x_1, \dots, x_n, x_\infty)$, called **$(n+2)$ -hyperspherical coordinates**. When $n=2$, they are called **tetracyclic coordinates**, and when $n=3$, **pentaspherical coordinates**. Therefore, if we restrict $(n+2)$ -hyperspherical coordinates for points on S^n , then they satisfy the quadratic relation stated before. In the frame $(A_0, A_1, \dots, A_n, A_\infty)$ of \mathbf{P}^{n+1} which defines the $(n+2)$ -hyperspherical coordinates, A_0 and A_∞ are points on S^n , and the other A_i are real hyperspheres passing through the points A_0 and A_∞ . It is possible to choose a frame $(A_0, A_1, \dots, A_n, A_\infty)$ such that the equation for S^n becomes $\sum_{i=1}^n x_i^2 - 2x_0 x_\infty = 0$ (i.e.,

$g_{ij} = \delta_{ij}$). Among hypersurfaces in S^n , one that is expressed by a homogeneous equation of the second degree with respect to $(n+2)$ -hyperspherical coordinates is called a **cyclide**. It is an algebraic surface of the fourth order and is an enveloping surface of the family of hyperspheres that are tangent to n fixed hyperspheres.

(6) Moving Coordinates. When we study the differential geometry of an m -dimensional surface W in a space M on which a transformation group G acts, it is often preferable to take a frame or frames at each point of W and consider a t connection among them. These frames are called **moving frames**, and the set of coordinate systems with respect to moving frames is called a **moving coordinate system** (\rightarrow 111 Differential Geometry of Curves and Surfaces).

C. Curvilinear Coordinates

Let (x_1, x_2, \dots, x_n) be a rectangular coordinate system on an n -dimensional Euclidean space E^n . If $x_i = x_i(u_1, u_2, \dots, u_n)$, $i = 1, \dots, n$, are functions of n variables (u_1, u_2, \dots, u_n) of class C^r ($r \geq 1$) and the t functional determinant $D(x_1, \dots, x_n)/D(u_1, \dots, u_n)$ is not equal to zero in some open domain, then (u_1, u_2, \dots, u_n) are considered local coordinates in E^n . We call them **curvilinear coordinates** of E^n . A hypersurface $u_i = \text{constant}$ (obtained by fixing the value of one of the variables u_i) is called a **coordinate hypersurface**, and a curve $u_j = \text{constant}$ ($j \neq i$) is called a **coordinate curve**. The line element ds of a Euclidean space E^n is given by

$$ds^2 = \sum_{k=1}^n dx_k^2 = \sum_{i,j=1}^n g_{ij} du_i du_j,$$

$$g_{ij} = \sum_{k=1}^n \frac{\partial x_k}{\partial u_i} \frac{\partial x_k}{\partial u_j}.$$

Thus E^n is equipped with a t Riemannian metric. However, as E^n is t flat, its t curvature tensor satisfies $R_{jkl}^i = 0$. If the metric is diagonal, namely, if $ds^2 = \sum_{i=1}^n g_i du_i^2$, the coordinates are called **orthogonal curvilinear coordinates**. Moreover, if $g_1 = \dots = g_n$, the coordinates are called **isothermal coordinates**. The metric is diagonal if and only if coordinate curves are mutually perpendicular at the points of intersection. Actually, curvilinear coordinate systems that are often used practically are diagonal. The concept of curvilinear coordinates has been generalized to the case of t differentiable manifolds and is utilized to determine their local coordinates.

On any 2-dimensional Riemannian manifold there always exist isothermal coordinates in a neighborhood of any point [7].

(1) Curvilinear Coordinates on Planes or Spaces (\rightarrow Appendix A, Table 3.V). Let (x, y) be rectangular coordinates of a point in a Euclidean plane E^2 . We have the following coordinate systems on E^2 :

Polar coordinates (r, θ) , where

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Elliptic coordinates (λ, μ) , where

$$x^2 = (\lambda + a^2)(\mu + a^2)/(a^2 - b^2),$$

$$y^2 = (\lambda + b^2)(\mu + b^2)/(b^2 - a^2),$$

$$a > b > 0, \quad \lambda > -b^2 > \mu > -a^2.$$

Parabolic coordinates (α, β) , where

$$x = -(\alpha + \beta), \quad y = \sqrt{-4\alpha\beta}, \quad \alpha > 0 > \beta.$$

Equilateral (or rectangular) hyperbolic coordinates (u, v) , where

$$x = uv, \quad y = (u^2 - v^2)/2.$$

Bipolar coordinates (ξ, η) , where

$$x = a \sin \xi / (\cosh \xi + \cos \eta),$$

$$y = a \sin \eta / (\cosh \xi + \cos \eta),$$

$$-\infty < \xi < \infty, \quad 0 \leq \eta \leq 2\pi.$$

Next we consider the case of a 3-dimensional Euclidean space E^3 and let (x, y, z) be rectangular coordinates on E^3 . The following systems of coordinates on E^3 are sometimes useful.

Cylindrical coordinates (r, θ, z) , where

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = z.$$

Spherical coordinates (r, θ, φ) , where

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta.$$

Ellipsoidal coordinates (λ, μ, ν) , where

$$x^2 = (\lambda + a^2)(\mu + a^2)(\nu + a^2)/(a^2 - b^2)(a^2 - c^2),$$

$$y^2 = (\lambda + b^2)(\mu + b^2)(\nu + b^2)/(b^2 - c^2)(b^2 - a^2),$$

$$z^2 = (\lambda + c^2)(\mu + c^2)(\nu + c^2)/(c^2 - a^2)(c^2 - b^2),$$

$$a > b > c > 0, \quad \lambda > -c^2 > \mu > -b^2 > \nu > -a^2.$$

These coordinate systems are all systems of orthogonal curvilinear coordinates. Suppose that we are given two rectangular coordinate systems (ξ, η, ζ) and (x, y, z) sharing the same origin. The correlation of the two is given by **Euler's angles** (θ, φ, ψ) , where θ , φ , and ψ are the angles between the z -axis and ζ -axis, zx -plane and $z\zeta$ -plane, and $\zeta\zeta$ -plane and ζz -plane, respectively. The Euler angles θ , φ , and ψ are subject to the inequalities $0 \leq \theta \leq \pi$ and $0 \leq \varphi, \psi < 2\pi$. They are often utilized in the dynamics of rigid bodies.

(2) Multipolar Coordinates. Let P_1, P_2, \dots, P_m be m points in general position in an n -dimensional Euclidean space E^n , $m \leq n$. If we

denote by $\rho_i (\geq 0)$ the distance between a point X of E^n and P_i , then $(\rho_1, \rho_2, \dots, \rho_m)$ can be regarded as coordinates of a point X contained in a suitable domain of E^n . They are called **multipolar coordinates**. In particular, if $m=2$ they are called **bipolar coordinates**, and if $m=3$, **tripolar coordinates**. When $m>n$, these coordinates satisfy $m-n$ relations. Next let $\alpha_1, \alpha_2, \dots, \alpha_m$ be m hyperplanes in general position in E^n , $m \leq n$. For an arbitrary point X of E^n , we denote by ξ_i the directed distance of X from each hyperplane α_i . The m -tuple $(\xi_1, \xi_2, \dots, \xi_m)$ provides coordinates of X that are called **multiplanar coordinates** in E^n . When $m>n$, these coordinates satisfy $m-n$ relations. In particular, when $n=2$ and $m=3$, they are called **trilinear coordinates**. In this case, if we denote by S the area of the triangle defined by three lines $\alpha_1, \alpha_2, \alpha_3$, and by a_1, a_2, a_3 the lengths of the three sides of the triangle, then the trilinear coordinates (ξ_1, ξ_2, ξ_3) satisfy a linear relation $a_1 \xi_1 + a_2 \xi_2 + a_3 \xi_3 = 2S$.

(3) Tangential Polar Coordinates. In a Euclidean plane E^2 , we take a directed line l_0 passing through a point O . For an arbitrary directed line g , let p be the directed distance between O and g , and let θ be the angle between l_0 and g . Then (p, θ) are called **tangential polar coordinates** (Fig. 1). They are useful for representing tangent lines to curves in E^2 . Let C be an \dagger oval in E^2 . A line is called a \dagger supporting line of C if its intersection with C consists of a point or a line segment. In this case we take the origin O inside C and consider the coordinates (p, θ) of the supporting lines of C . Then the equation of C can be represented as $p = p(\theta)$, where $p(\theta)$ is a periodic function of period 2π . The coordinates (p, θ) are especially useful when the function $p(\theta)$ can be expanded in a \dagger Fourier series. In the case of Euclidean space E^3 , the notion of tangential polar coordinate system can also be defined by using tangent planes.

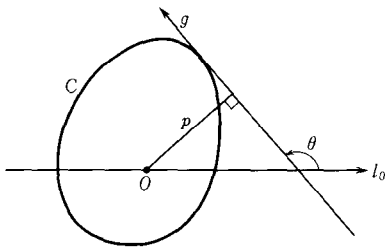


Fig. 1
Tangential polar coordinates.

(4) Normal Coordinates. Let M be an n -dimensional \dagger Riemannian manifold, and let $T_A(M)$ be the tangent space to M at a point A . For each tangent vector $\mathbf{v} \in T_A(M)$, we draw a

\dagger geodesic through A with the initial direction \mathbf{v} and take a point P on the geodesic such that the distance from A to P is equal to the length of \mathbf{v} . Then the correspondence that sends \mathbf{v} to P is a \dagger diffeomorphism of a neighborhood of the zero vector 0 of $T_A(M)$ with a neighborhood of A in M . Therefore the components (v^1, v^2, \dots, v^n) of \mathbf{v} with respect to a basis of $T_A(M)$ give the coordinates of the points P contained in a suitable neighborhood of A . We call them **normal coordinates** about the point A of M . In these coordinates, each geodesic passing through A is given by equations $v^i = \alpha^i r$ ($i=1, 2, \dots, n$), where the (α^i) are components of the unit vector in the direction of \mathbf{v} and r is the parameter that represents the arc length from A to the point (v^1, \dots, v^n) . In particular, when $n=2$, we fix a tangent vector \mathbf{v}_0 at A and denote the angle between \mathbf{v} and \mathbf{v}_0 by θ . Then (r, θ) are coordinates of P called **geodesic polar coordinates**. The notion of normal coordinates can also be defined for \dagger Lie groups or differentiable manifolds with \dagger affine connections.

D: Local Coordinates

Suppose that we have a space M that has a covering by a family of open neighborhoods with coordinate systems. If, for each pair of neighborhoods with nonempty intersection, the coordinate transformation in the intersection satisfies certain specified conditions, then a mathematical structure on M can be defined.

Let E be a \dagger topological space. Suppose that Φ is a family of open sets in E such that the union of any number of open sets in Φ and the intersection of any finite number of open sets in Φ also belong to Φ . A set Γ of \dagger homeomorphisms is called a **pseudogroup of transformations** on E if Γ satisfies the following three conditions: (i) Any homeomorphism $f \in \Gamma$ is defined on an open set $U \in \Phi$, and $f(U) \in \Phi$. (ii) When an open set $U \in \Phi$ is expressed as the union of a family $\{U_i\}$ of open sets $U_i \in \Phi$, a homeomorphism f defined on U belongs to Γ if and only if its restriction to each U_i belongs to Γ . (iii) For any open set $U \in \Phi$, the identity mapping on U belongs to Γ , and if $f, g \in \Gamma$, then the inverse f^{-1} and the composition $g \circ f$, if it exists, belong to Γ .

Let E and M be topological spaces. A homeomorphism $\varphi: U \rightarrow V$ of an open set U in E to an open set V in M is called a **local coordinate system** of M with respect to E . For two local coordinate systems $\varphi_1: U_1 \rightarrow V_1$ and $\varphi_2: U_2 \rightarrow V_2$, the homeomorphism

$$\varphi_2^{-1} \circ \varphi_1: \varphi_1^{-1}(V_1 \cap V_2) \rightarrow \varphi_2^{-1}(V_1 \cap V_2)$$

is called a **transformation of local coordinates**.

Let Γ be a pseudogroup of transformations

on E . A set Σ of local coordinate systems of M with respect to E is said to define a Γ -structure on M if Σ satisfies the following two conditions: (1) the totality of the images of local coordinate systems belonging to Σ covers M ; (2) if two local coordinate systems ϕ_1 and ϕ_2 of Σ have a transformation of local coordinates, it belongs to Γ . Now suppose that two sets Σ and Σ' of local coordinate systems define Γ -structures on M . If the union of Σ and Σ' defines a Γ -structure, then we say that the first two Γ -structures are equivalent. Let Γ be a pseudogroup of t diffeomorphisms, each defined from an open subset of the n -dimensional space \mathbf{R}^n onto another open set. If a Γ -structure is defined on a space M , then M is an n -dimensional t differentiable manifold.

On the other hand, let Γ be a pseudogroup of complex analytic homeomorphisms in an n -dimensional complex number space \mathbf{C}^n . If a Γ -structure is defined on a space M , then M is an n -dimensional t complex analytic manifold. Locally homogeneous spaces, t foliated manifolds, the t fiber bundles are all equipped with local coordinate systems with suitable Γ -structures.

References

- [1] E. Cartan, La méthode du repère mobile, la théorie des groupes continus et les espaces généralisés, Actualités Sci. Ind., Hermann, 1935.
- [2] C. Ehresmann, Sur la théorie des espaces fibrés, Colloques Intern. Topologie algébrique, Paris, 1949, C. N. R. S. XII, 3–15.
- [3] F. Klein, Vorlesungen über höhere Geometrie, Springer, third edition, 1926 (Chelsea, 1957).
- [4] B. L. van der Waerden, Einführung in die algebraische Geometrie, Springer, 1939.
- [5] O. Veblen and J. H. C. Whitehead, The foundations of differential geometry, Cambridge Univ. Press, 1932.
- [6] Wu Wen-Tsun and G. Reeb, Sur les espaces fibrés et les variétés feuilletées, Actualités Sci. Ind., Hermann, 1952.
- [7] S. S. Chern, An elementary proof of the existence of isothermal parameters on a surface, Proc. Amer. Math. Soc., 6 (1955), 771–782.

91 (IX.3) Covering Spaces

A. General Remarks

A continuous mapping $p: \tilde{Y} \rightarrow Y$ of an t arcwise connected topological space \tilde{Y} onto a con-

nected topological space Y is called a **covering mapping (covering map)** if the following condition (C) is satisfied: (C) Each point of Y has an open neighborhood V such that every t connected component of $p^{-1}(V)$ is mapped homeomorphically onto V by p . Here we note that Y is in fact arcwise connected.

If there is a covering mapping $p: \tilde{Y} \rightarrow Y$, we call \tilde{Y} a **covering space** of Y and (\tilde{Y}, p, Y) a **covering**. In particular, for a t differentiable manifold Y , if \tilde{Y} is also a differentiable manifold and p is differentiable, then \tilde{Y} is called a **covering (differentiable) manifold** of Y . (In the theory of t Riemann surfaces, a t covering surface may have some t branch points violating condition (C). Upon removing such points, we obtain a covering space as defined above.)

For each t path $w: I \rightarrow Y$ ($I = [0, 1]$) of Y , a path $\tilde{w}: I \rightarrow \tilde{Y}$ with $p \circ \tilde{w} = w$ is uniquely determined by the point $\tilde{w}(1) \in p^{-1}(w(1))$, and a bijection $w_{\#}: p^{-1}(w(1)) \rightarrow p^{-1}(w(0))$ is determined by $w_{\#}(\tilde{w}(1)) = \tilde{w}(0)$. Thus there exists a one-to-one correspondence between $p^{-1}(y)$ and $p^{-1}(y')$ for every pair of points y, y' of Y , and $(\tilde{Y}, p, Y, p^{-1}(y_0))$ is a t locally trivial fiber space with discrete fiber $p^{-1}(y_0)$. When the cardinal number of $p^{-1}(y)$ is a finite number n , we call (\tilde{Y}, p, Y) an n -fold covering. In this case, for a t loop $w(I, \dot{I}) \rightarrow (Y, y_0)$ with base point y_0 , $w_{\#}: p^{-1}(y_0) \rightarrow p^{-1}(y_0)$ is a permutation of the n elements in $p^{-1}(y_0)$, and we obtain a homomorphism of the t fundamental group $\pi_1(Y) = \pi_1(Y, y_0)$ of Y into the t symmetric group \mathfrak{S}_n , given by the correspondence $w \rightarrow w_{\#}$. The permutation group \mathfrak{M} , which is the image of this homomorphism, is called the **monodromy group** of the n -fold covering.

Two coverings (\tilde{Y}_i, p_i, Y) ($i = 1, 2$) are said to be **equivalent** if there is a homeomorphism $\phi: \tilde{Y}_1 \rightarrow \tilde{Y}_2$ with $p_2 \circ \phi = p_1$; such a ϕ is called an **equivalence**. In particular, a self-equivalence $\phi: \tilde{Y} \rightarrow \tilde{Y}$ of a covering (\tilde{Y}, p, Y) is called a **covering transformation**. The set π of all covering transformations forms a group by the composition of mappings, which is called the **covering transformation group** of \tilde{Y} . We call (\tilde{Y}, p, Y) a **regular covering** if for each $y \in Y$ and $\tilde{y}_1, \tilde{y}_2 \in p^{-1}(y)$, there exists a unique covering transformation that maps \tilde{y}_1 to \tilde{y}_2 . In this case, the t orbit space \tilde{Y}/π is homeomorphic to Y , (\tilde{Y}, p, Y, π) is a t principal bundle, and the monodromy group \mathfrak{M} is isomorphic to π .

For a covering (\tilde{Y}, p, Y) , we call \tilde{Y} a **covering group** of Y if \tilde{Y} and Y are topological groups and p is a homomorphism. Then (\tilde{Y}, p, Y) is a regular covering, and its covering transformation group is isomorphic to $p^{-1}(e)$ (e is the identity element of Y), which is a discrete subgroup lying in the center of \tilde{Y} (\rightarrow 423 Topological Groups O).

B. Universal Covering Spaces

For a covering (\tilde{Y}, p, Y) , we have the following relations of the \dagger homotopy groups: $p_*: \pi_i(\tilde{Y}) \rightarrow \pi_i(Y)$ is isomorphic (monomorphic) for $i \geq 2$ ($i = 1$), and $\pi_1(Y)/p_*(\pi_1(\tilde{Y}))$ is in one-to-one correspondence with $p^{-1}(y_0)$ ($y_0 \in Y$). If \tilde{Y} is \dagger simply connected, \tilde{Y} is called a **universal covering space** of Y ; if in addition \tilde{Y} is a covering group, \tilde{Y} is called a **universal covering group** of Y .

For a \dagger locally arcwise connected space Y , a covering (\tilde{Y}, p, Y) is regular if and only if $p_*(\pi_1(\tilde{Y}))$ is a normal subgroup of $\pi_1(Y)$, and a universal covering space of Y is a covering space of any covering space of Y . Moreover, if Y is a topological group, any covering space \tilde{Y} of Y can be given a unique topological group structure with which \tilde{Y} is a covering group of Y .

Let Y be an arcwise connected, locally arcwise connected, and \dagger locally simply connected space. Then the following classification theorem of coverings holds: The set of equivalence classes of coverings of Y is in one-to-one correspondence with the set of conjugate classes of subgroups of the fundamental group $\pi_1(Y)$; in particular, the equivalence class of a covering (\tilde{Y}, p, Y) corresponds to the conjugate class of the subgroup $p_*(\pi_1(\tilde{Y}))$. Also, there is a unique universal covering space \tilde{Y} of Y up to homeomorphism. If in addition Y is a topological group, then \tilde{Y} is a unique universal covering group of Y up to isomorphism of topological groups. Such a space \tilde{Y} is obtained as follows: Consider the \dagger path space $\Omega(Y; y_0, Y)$ of all paths in Y starting from a fixed point $y_0 \in Y$, and define two paths $w_0, w_1: (I, 0) \rightarrow (Y, y_0)$ such that $w_0(1) = w_1(1)$ to be equivalent if and only if there is a \dagger homotopy $w_t: (I, 0) \rightarrow (Y, y_0)$ with $w_t(1) = w_0(1)$ ($0 \leq t \leq 1$). Then we obtain the \dagger identification space \tilde{Y} of $\Omega(Y; y_0, Y)$ by this equivalence relation and the mapping $p: \tilde{Y} \rightarrow Y$ by $p\{w\} = w(1)$; this \tilde{Y} is the universal covering space of Y .

Let (\tilde{Y}, p, Y) be a regular covering with the covering transformation group π . Then there is a \dagger locally trivial fiber space (Y', q, B, \tilde{Y}) such that the total space Y' has the same \dagger (co) homology groups as Y , and the base space B is an \dagger Eilenberg-MacLane space $K(\pi, 1)$. The (co) homology \dagger spectral sequence of this fiber space is called that of the given regular covering (\tilde{Y}, p, Y) , E_∞ is a bigraded module \dagger associated with a certain filtration of the \dagger singular (co) homology module $H(Y)$, and E_2 is the (co) homology module $H(\pi; H(\tilde{Y}))$ of the group π , where π operates on the coefficient module $H(\tilde{Y})$ via the induced homomorphisms of covering transformations (\rightarrow 148 Fiber Spaces).

For any group π , there is a regular covering (E_π, p, B_π) with the covering transformation group π such that E_π is \dagger contractible; B_π is an Eilenberg-MacLane space $K(\pi, 1)$. We can take S^1 (1-sphere) as $B_{\mathbb{Z}}$ of the infinite cyclic group \mathbb{Z} , and the following infinite lens space as $B_{\mathbb{Z}_k}$ of the finite cyclic group \mathbb{Z}_k (\rightarrow 70 Complexes).

C. Lens Spaces

Let k be a positive integer and l_1, \dots, l_n be integers prime to k . Let $S^{2n+1} = \{(z_0, \dots, z_n) \in \mathbb{C}^{n+1} \mid |z_0|^2 + \dots + |z_n|^2 = 1\}$ be the unit sphere in the $(n+1)$ -dimensional complex linear space \mathbb{C}^{n+1} , and define the rotation γ by $\gamma(z_0, z_1, \dots, z_n) = (z_0 \exp 2\pi i/k, z_1 \exp 2\pi i l_1/k, \dots, z_n \exp 2\pi i l_n/k)$. Then the \dagger orbit space $S^{2n+1}/\mathbb{Z}_k = L(k; l_1, \dots, l_n)$, where $\mathbb{Z}_k = \mathbb{Z}/k\mathbb{Z}$ is interpreted as the cyclic group generated by γ , is called a **lens space**. It is an orientable $(2n+1)$ -dimensional \dagger differentiable manifold. Also, the **infinite lens space** $L^\infty(k) = L(k; 1, \dots, 1, \dots)$ is defined by taking $n = \infty$; the infinite sphere S^∞ is a k -fold covering space of $L^\infty(k)$, and $L^\infty(k) = B_{\mathbb{Z}_k} = K(\mathbb{Z}_k, 1)$. Its \dagger cohomology ring is given as follows.

(1) For integral coefficients, $H^{2i+1}(L^\infty(k)) = 0$, $H^{2i}(L^\infty(k)) = \mathbb{Z}_k$ ($i > 0$), and the \dagger cup product of generators of degree $2i$ and $2j$ is a generator of degree $2(i+j)$.

(2) Let $k = p^k$ (p is a prime). If $p \neq 2$, or $p = 2$ and k' is even, $H^*(L^\infty(k); \mathbb{Z}_p) = \wedge(e_1) \otimes \mathbb{Z}_p[e_2]$. If $p = 2$ and k' is odd, $H^*(L^\infty(k); \mathbb{Z}_2) = \mathbb{Z}_2[e_1]$ (e_i is an element of degree i). Here \wedge indicates the \dagger exterior algebra over \mathbb{Z}_p , and $\mathbb{Z}_p[\]$ the \dagger polynomial ring over \mathbb{Z}_p .

Two lens spaces $L(k; l_1, \dots, l_n)$ and $L(k'; l'_1, \dots, l'_n)$ are of the same homotopy type if and only if $k = k'$ and there is an integer m prime to k with

$$l_1 \dots l_n \equiv \pm m^{n+1} l'_1 \dots l'_n \pmod{k}$$

[8]. Furthermore, the condition $k = k'$, $l \equiv \pm l' \pmod{k}$ holds if and only if $L(k; l)$ and $L(k'; l')$ are homeomorphic. (Sufficiency is shown in [1]; necessity follows from the fact that the \dagger Hauptvermutung is valid for combinatorial 3-manifolds and that the condition holds if the polyhedra $L(k; l)$ and $L(k'; l')$ have isomorphic subdivisions [6].) Also \rightarrow [7, 11] and 65 Combinatorial Manifolds.

References

- [1] H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Teubner, 1934 (Chelsea, 1965).
- [2] N. Steenrod, *The topology of fiber bundles*, Princeton Univ. Press, 1951.
- [3] S.-T. Hu, *Homotopy theory*, Academic Press, 1959.

- [4] S. Eilenberg and S. MacLane, Homology of spaces with operators II, Trans. Amer. Math. Soc., 65 (1949), 49–99.
- [5] J. H. C. Whitehead, On incidence matrices, nuclei and homotopy types, Ann. Math., (2) 42 (1941), 1197–1239.
- [6] K. Reidemeister, Homotopieringe und Linsenräume, Abh. Math. Sem. Univ. Hamburg, 11 (1935), 102–109.
- [7] W. Franz, Über die Torsion einer Überdeckung, J. Reine Angew. Math., 173 (1935), 245–254.
- [8] P. Olum, Mappings of manifolds and the notion of degree, Ann. Math., (2) 58 (1953), 458–480.
- [9] W. S. Massey, Algebraic topology; An introduction, Harcourt, Brace & World, 1967.
- [10] I. M. Singer and J. A. Thorpe, Lecture notes on elementary topology and geometry, Scott, Foresman, 1967.
- [11] M. M. Cohen, A course in simple homotopy theory, Springer, 1972.

92 (IV.15) Crystallographic Groups

A. Space Groups and Point Groups

Let G be a discrete subgroup of the group of motions \mathfrak{G} in the real n -dimensional Euclidean space V . If G contains n linearly independent translations, then G is called an n -dimensional **crystallographic space group**, a **space group**, or a **crystallographic group**. The following assertions are known to be equivalent to each other: (i) A discrete subgroup G of \mathfrak{G} is a space group; (ii) \mathfrak{G}/G is compact; (iii) V/G is compact; (iv) There is a compact subset P of V such that $V = GP = \bigcup_{g \in G} gP$ [1]. The subgroup T consisting of all translations in G is called the **lattice group** of G . T is a normal subgroup of G and is generated by n linearly independent translations, say, t_1, t_2, \dots, t_n . Take a point x of V . Then the T -orbit of the point x is called a **lattice** of G (or T). t_1, \dots, t_n can be identified with a basis of V . Let K be the quotient group G/T and $k = gT$ an element of K . Then k gives rise to a linear transformation \bar{k} on T via the formula $\bar{k}(t) = gtg^{-1}$ (the product in G), $t \in T$, and K can be regarded as a subgroup of the orthogonal group of V via the above identification of T with a lattice in V . That is, the set of orthogonal transformations stabilizing x , $\{\bar{g} | (\bar{g}(y) = x + g(y) - g(x), y \in V), g \in G\}$ forms a group isomorphic to K . The quotient group K is a finite group and is called the **point group** of G . A group isomorphic to a point group of an n -dimensional space group is sometimes called

an n -dimensional **crystallographic group**. An algebraic characterization of space groups among abstract groups is given by “A group G is isomorphic to an n -dimensional space group $\Leftrightarrow G$ has a normal free Abelian subgroup of rank n , which is maximal Abelian and has finite index” [1].

By representing the action of K on T in terms of a basis of T , K can be regarded as a subgroup of $GL(n, \mathbb{Z})$. Hence we obtain the following **crystallographic restriction**: If K contains an element of order m , then $n \geq \varphi(m)$. Here φ is the Euler function. For example, if $n = 2, 3$, then $m = 1, 2, 3, 4$ or 6 ; and if $n = 4, 5$, then $m = 1, 2, 3, 4, 5, 6, 8, 10$, or 12 . The number of nonconjugate finite subgroups of $GL(n, \mathbb{Z})$ is finite (Jordan-Zassenhaus theorem; $\rightarrow [2, 3, 21]$).

Two space groups G_1 and G_2 are called **equivalent** if they are conjugate via an affine transformation f of V , i.e., $G_2 = fG_1f^{-1}$. Let T_i and K_i be the lattice group and the point group of G_i ($i = 1, 2$), respectively. Then $T_2 = hT_1$ and $K_2 = hK_1h^{-1}$ for the linear transformation h induced by f . Moreover, there is an orthogonal transformation k such that $K_2 = kK_1k^{-1}$; this k can naturally be considered as an element of $GL(n, \mathbb{Z})$. G_1 and G_2 are equivalent $\Leftrightarrow G_1$ and G_2 are isomorphic (as abstract groups) [1]. In applications it is often required that in the definition of equivalence f be an orientation-preserving affine transformation. Under this definition, there can be a pair G_1, G_2 such that G_1 and G_2 are mutually isomorphic but not equivalent, since they are conjugate only by means of an orientation-reversing affine transformation. In this case, they are called **enantiomorphous** to each other, or the pair is called an **enantiomorphic pair** (Table 1 below). For a given dimension n , there is only a finite number of equivalence classes of space groups [1]. We also conclude from this that point groups have finitely many equivalence classes up to conjugacy in the orthogonal group $O(V)$ or in $GL(n, \mathbb{Z})$.

B. Crystal Classes, Bravais Types

Let T be an n -dimensional lattice in n -dimensional Euclidean space V , and K a finite subgroup of the orthogonal group $O(V)$. Denote by a pair (T, K) a faithful linear representation of K on T , i.e., a monomorphism of groups $K \rightarrow \text{Aut}(T)$. A space group G determines a pair $(T, G/T)$. Conversely, for any pair (T, K) the following holds: “Any group extension of K over the kernel T is isomorphic to a space group” (\rightarrow the algebraic characterizations in Sections A and C). Thus to each pair (T, K) there corresponds a set of space groups.

Two pairs (T_1, K_1) and (T_2, K_2) are called **arithmetically equivalent** (this is denoted by $(T_1, K_1) \approx (T_2, K_2)$) if there exists an invertible linear transformation $g \in GL(V)$ such that $T_2 = gT_1$, $K_2 = gK_1g^{-1}$. These pairs are called **geometrically equivalent** or simply **equivalent** (this is denoted by $(T_1, K_1) \sim (T_2, K_2)$) if there exists a $g \in GL(V)$ such that $K_2 = gK_1g^{-1}$. The relation $\approx (\sim)$ is an equivalence relation, and equivalence classes are called **arithmetic (geometric) crystal classes**. The set of geometric crystal classes is in one-to-one correspondence with the set of equivalence classes of point groups up to conjugacy, and is sometimes called the set of **crystal classes**. The geometric crystal class of (T, K) is usually denoted by K .

Now let T be a lattice. Then the group of all orthogonal transformations that leave T invariant is called the **Bravais group** of T and is denoted by $B(T)$, i.e., $B(T) = \{g \in O(V) | gT = T\}$. The group $B(T)$ is finite and determines a pair $(T, B(T))$. Two lattices T_1 and T_2 are called **arithmetically equivalent** (this is denoted by $T_1 \approx T_2$) if $(T_1, B(T_1))$ and $(T_2, B(T_2))$ are arithmetically equivalent. An equivalence class is called a **Bravais type**, and an arithmetic crystal class determined by $(T, B(T))$ is called a **Bravais class**. For each Bravais type, a representative is said to be its **Bravais lattice** (Fig. 3 below).

Define an \dagger order relation on the set of lattices belonging to an arithmetic crystal class as follows. When $(T_1, K_1) \approx (T_2, K_2)$, define $T_2 \leq T_1$ if there exists a $g \in GL(V)$ such that $T_2 = gT_1$, $K_2 = gK_1g^{-1}$, and $B(T_2) \subseteq gB(T_1)g^{-1}$. Note that $T_1 \approx T_2 \Leftrightarrow T_1 \leq T_2$ and $T_2 \leq T_1$. In each arithmetic crystal class, there is a pair (T, K) with a \dagger minimal T with respect to this order relation. In this case the Bravais type of T is referred to as the **Bravais type** of the class of (T, K) . This lattice T is, intuitively, of the most general type appearing in the class. Now define a relation $T_1 \sim_f T_2$ if there are pairs (T_i, K_i) , $i = 1, 2$, such that $(T_1, K_1) \sim (T_2, K_2)$ and T_i is minimal in the class of (T_i, K_i) . The equivalence class of the equivalence relation generated by the relation \sim_f is called a **crystal family**.

Summarizing, we have defined the following. Let $\mathcal{S}, \mathcal{A}, \mathcal{G}, \mathcal{B}, \mathcal{CF}$ be the set of equivalence classes of space groups, arithmetic crystal classes, geometric crystal classes, Bravais types, and crystal families, respectively. Then we have the relations shown in Fig. 1. (All arrows in Fig. 1 are surjective.) \mathcal{CF} is nothing but the \dagger coproduct of \mathcal{G} and \mathcal{B} over \mathcal{A} , $\mathcal{CF} \simeq \mathcal{G} \amalg_{\mathcal{A}} \mathcal{B}$.

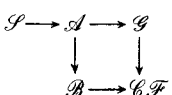


Fig. 1

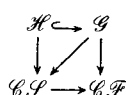


Fig. 2

Moreover, there are injective mappings $\mathcal{B} \hookrightarrow \mathcal{A}$ and $\mathcal{A} \hookrightarrow \mathcal{S}$ such that the composites $\mathcal{B} \hookrightarrow \mathcal{A} \rightarrow \mathcal{B}$ and $\mathcal{A} \hookrightarrow \mathcal{S} \rightarrow \mathcal{A}$ are identities. The first is the mapping sending a Bravais type T to the arithmetic crystal class defined by $(T, B(T))$. The second one sends a class of (T, K) to a \dagger semidirect product of T and K (\rightarrow 190 Groups N). A space group belonging to a class of the image of $\mathcal{A} \hookrightarrow \mathcal{S}$ is said to be **symmorphic** or **symmorphous**. A class which belongs to the image \mathcal{H} of the composite $\mathcal{B} \hookrightarrow \mathcal{A} \rightarrow \mathcal{G}$ is called a **holohedral** or **holosymmetric class**, or a **holohedry**. Define a mapping M of \mathcal{G} to the power set $\mathfrak{P}(\mathcal{B})$ by $M = (\mathcal{B} \leftarrow \mathcal{A}) \cdot (\mathcal{A} \rightarrow \mathcal{G})^{-1}$. For two elements C_1 and C_2 of \mathcal{G} , define $C_1 \sim_s C_2$ if $M(C_1) = M(C_2)$. Each equivalence class is called a **crystal system**, and the set of crystal systems is denoted by \mathcal{CS} . Two elements C and C' of \mathcal{G} belong to the same crystal family if and only if there exists a sequence of elements of \mathcal{G} , C_1, C_2, \dots, C_k such that $C_1 = C$ and $C_k = C'$ and $M(C_i) \cap M(C_{i+1}) \neq \emptyset$ for $i = 1, \dots, k-1$. Therefore there is a surjective mapping $\mathcal{CS} \rightarrow \mathcal{CF}$ (Fig. 2). When $n \leq 4$, the composite $\mathcal{H} \hookrightarrow \mathcal{G} \rightarrow \mathcal{CS}$ is bijective. It should be remarked that when $n \geq 3$, there is no mapping $\mathcal{B} \rightarrow \mathcal{CS}$ such that $\mathcal{A} \rightarrow \mathcal{G} \rightarrow \mathcal{CS}$ and $\mathcal{A} \rightarrow \mathcal{B} \rightarrow \mathcal{CS}$ coincide.

The numbers of elements of these sets are shown in Table 1.

Table 1

	$n =$	1	2	3	4
\mathcal{CF} , crystal families		1	4	6	23
\mathcal{CS} , ^a crystal systems		1	4	7	33
\mathcal{B} , Bravais types		1	5	14	64
\mathcal{G} , point groups (isomorphic classes) ^b		2 (2)	10 (9)	32 (18)	227 (118)
\mathcal{A} , arithmetic crystal classes		2	13	73	710
\mathcal{S} , space groups		2	17	219 ^c	4783 ^d

^a When $n \leq 4$, the number of elements of \mathcal{H} coincides with that of \mathcal{CS} .

^b The number in parentheses denotes the number of equivalence classes under algebraic isomorphism.

^{c,d} The number of equivalence classes under orientation-preserving transformations is 230 and 4895, respectively.

C. Construction of Space Groups

Take an element (T, K) of an arithmetic crystal class. Since K is finite and T is \dagger finitely generated, the \dagger cohomology groups $H^i(K, T)$ are finite groups (\rightarrow 200 Homological Algebra G). Let α be a \dagger second cocycle representing an element $[\alpha]$ of $H^2(K, T)$. Let $(T, K)_\alpha$ be the \dagger extension of K with kernel T corresponding to $[\alpha]$ (\rightarrow 190 Groups N). Then the set of elements of $(T, K)_\alpha$ is given by $\{\langle t, k \rangle |$

$t \in T, k \in K$, and the product is defined by $\langle t_1, k_1 \rangle \langle t_2, k_2 \rangle = \langle t_1 + k_1 t_2 + \alpha(k_1, k_2), k_1 k_2 \rangle$. Since $H^1(K, V) = 0, i \geq 1$, there is a \dagger first co-chain $\beta \in C^1(K, V)$ such that $\alpha(k_1, k_2) = \beta(k_1) + k_1 \beta(k_2) - \beta(k_1 k_2)$, and the \dagger first cocycle β on V/T defined by β induces an isomorphism $H^1(K, V/T) \cong H^2(K, T), [\beta] \mapsto [\alpha]$. Define an action of $\langle t, k \rangle$ on V by

$$\langle t, k \rangle(x) = kx + t + \beta(k), \quad x \in V.$$

Then $(T, K)_\alpha$ acts on V as a space group. Concerning their equivalence, this result holds:

"The equivalence classes of $(T, K)_\alpha, [\alpha] \in H^2(K, T)$ are in one-to-one correspondence with the orbits of $N(K)$, the normalizer of K in $\text{Aut}(T)$ acting on $H^2(K, T)$ " [4]. The action of $h \in N(K)$ is given by ${}^h\alpha(k) = h\alpha(h^{-1}kh), \alpha \in C^2(K, T)$.

For a pair (T, K) the **weight lattice** or the **weight group** T^* of (T, K) is defined by $\{v \in V \mid v - kv \in T, \forall k \in K\}$. Let R be an irreducible reduced \dagger root system, $Q(R)$ its lattice, and $W(R)$ its \dagger Weyl group. Then for a pair $(Q(R), W(R))$, $Q(R)^*$ is nothing but the usual weight lattice $P(R)$. (\rightarrow 248 Lie Algebras). Suppose that K of (T, K) does not contain the central inversion $-I_n: v \mapsto -v, v \in V$. Denote the group $K \cup (-I_n)K$ by $\pm K$. Then there is an isomorphism

$$H^1(\pm K, V/T) \cong [H^1(K, V/T)]_2 \times (T^*/T)/2(T^*/T).$$

Here $[\cdot]_2$ denotes the subgroup of elements of order not greater than 2 [5]. This isomorphism provides powerful means for the construction of space groups.

In constructing a pair (T, K) , it is useful to take a basis of T in a special form. Let $\{t_1, t_2, \dots, t_n\}$ be a basis of a lattice T , and set $a_{ij} = (t_i, t_j)$, where (\cdot) is the inner product in V . The $n \times n$ matrix $A = (a_{ij})$ is \dagger symmetric and \dagger positive definite, and the numbers a_{ij} are called the **lattice constants** of T . A basis $\{t_1, \dots, t_n\}$ is called a **reduced basis** if the \dagger quadratic form $(Ax, x) = \sum a_{ij}x_i x_j$ is reduced (\rightarrow 348 Quadratic Forms). A lattice admits at least one reduced basis, and in this case $\pm 2a_{ij} \leq a_{ii} \leq a_{jj}, i < j$ [6, Table 5.1; 7]. The totality of an n -dimensional reduced basis forms in a natural manner a semialgebraic set of dimension $(n+2)(n-1)/2$ in $\mathbf{R}^{(n+1)n/2}$. Its closure, say Q_n , is compact. On the other hand, the set $L_n = (\mathbf{R}_+ \times O(\mathbf{R}^n)) \backslash GL(n, \mathbf{R})/GL(n, \mathbf{Z})$ is identified with the set of equivalence classes of n -dimensional lattices under the relation of "same shape." There is a natural inclusion $L_1 \cup \dots \cup L_n \subset Q_n$, and L_n is dense in Q_n . Moreover, the inclusion gives a bijective homeomorphism when $n \leq 4$ [8].

A finite subgroup K of the orthogonal group $O(V)$ is said to be **fully transitive** if there

is a set $S = \{e_1, \dots, e_s\}$ that spans V on which K acts \dagger transitively and if K has no \dagger invariant subspace in V . In this case one can choose S as either of the following: (i) the **primitive hypercubic** type, $S = \{e_1, \dots, e_n\}, (e_i, e_j) = \delta_{ij}$ (\dagger Kronecker delta); (ii) the **primitive hyperbolic** type, $S = \{f_1, \dots, f_{n+1}\}, (f_i, f_i) = 1, i = 1, \dots, n+1, (f_i, f_j) = -1/n, i, j = 1, \dots, n+1, i \neq j$ (especially, $\sum_{i=1}^{n+1} f_i = 0$) [10].

D. Color Symmetry Groups, Twinning

A **color symmetry group**, or a **colored symmetry group**, associated with a space group G is a pair (G, G') such that G' is a subgroup of G with finite index. The index $r = [G : G']$ is called the **number of colors**. We call (G, G') a **white group** when $r = 1$, a **black and white group** or a **magnetic group** when $r = 2$, and a **polychromatic group** when $r \geq 3$. G' is also a space group with the lattice group $T' = G' \cap T$, where T is the lattice group of G . Two color symmetry groups (G_1, G'_1) and (G_2, G'_2) are called **equivalent** if there exists an \dagger affine transformation f such that $G_2 = fG_1f^{-1}$ and $G'_2 = fG'_1f^{-1}$. Then (G_1, G'_1) and (G_2, G'_2) are equivalent if and only if there is an isomorphism $\sigma: G_1 \rightarrow G_2$ such that $\sigma(G'_1) = G'_2$. A color symmetry group (G, G') is called **lattice equivalent** if $T' = T, K' \subset K$, and is called **class equivalent** if $T' \subset T, K' = K$. Take a pair (G, G') , and let G'' be the \dagger inverse image of K' under the canonical \dagger epimorphism $G \rightarrow K$. Then (G, G'') is lattice equivalent and (G'', G') is class equivalent. Fix the color number r . Then (i) the equivalence classes of lattice equivalent color symmetry groups are in one-to-one correspondence with the conjugacy classes of subgroups K_1 of index r , and (ii) the number of equivalence classes of class equivalent color symmetry groups is finite. When r is prime to the order of K , the classes are in one-to-one correspondence with K -invariant sublattices T_1 of index r in T . Therefore, when the number of colors is given, the set of equivalence classes of n -dimensional color symmetry groups is finite.

In case (i), a pair (K, K_1) is called a **color point group**, and in case (ii), (T, T_1) is called a **color lattice**. In particular, when $r = 2$, they are called a **black and white point group** and a **black and white lattice**, respectively. Their equivalence is defined in a similar way.

Take a color lattice (T, T_1) of color number r . Let m be the least natural number such that $mT \subset T_1$, then m is a divisor of r . Let $m = p_1^{n_1} \dots p_k^{n_k}, p_i \neq p_j (i \neq j)$ be the decomposition of m into its prime factors. Then there is a sequence of K -invariant sublattices of $T, T_i, i = 2, \dots, k-1$, such that (i) $T_1 \subset T_2 \dots \subset T_k \subset T_{k+1} = T$; (ii) $p_i^{n_i} T_{i+1} \subset T_i, i = 1, \dots, k$; (iii) the

submodules $T_i/p_i^{n_i}T_{i+1}$ of $i = 1, \dots, k$, $\mathbf{Z}^n \otimes \mathbf{Z}/p_i^{n_i}\mathbf{Z}$ are uniquely determined independently of the choice of such a sequence $\{T_i\}$. Thus the existence of color lattices is related to the \dagger modular representation of K (\rightarrow 362 Representations G). For example, if r is a prime number, then the representation of K on $T/rT \simeq \mathbf{Z}^n \otimes \mathbf{Z}/r\mathbf{Z}$ is the reduction modulo r of that of K on T .

A mathematical treatment of twinning is given as follows. Let V be an n -dimensional Euclidean space and T an n -dimensional lattice in V . The volume of T , $\text{vol}(T)$, is defined by $\text{vol}(T) = (|\det A|)^{1/2}$, where $A = (a_{ij})$ is the matrix of lattice constants. This definition is independent of the choice of a basis of T . If T' is a sublattice of index r in T , $\text{vol}(T') = r \cdot \text{vol}(T)$ follows. Let K be a finite subgroup of $O(V)$, T_1 and T_2 K -invariant lattices, and U a hyperplane in V . Then the quadruple (T_1, T_2, K, U) is called a **twinning structure** if it satisfies the following: (i) $\text{vol}(T_1) = \text{vol}(T_2)$; (ii) the lattice T_i is minimal in the arithmetic crystal class of (T_i, K) , $i = 1, 2$; and (iii) $T_1 \cap U = T_2 \cap U = T_0$, and T_0 is an $(n-1)$ -dimensional lattice in U . Two arithmetic crystal classes are said to be **twinable** if there exist representatives (T_1, K) , (T_2, K) , and a hyperplane U such that (T_1, T_2, K, U) is a twinning structure. Let (T_1, K) and (T_2, K) be two arithmetic crystal classes. If there exists a K -invariant hyper-surface U such that $T_0 = T_1 \cap U = T_2 \cap U$ is an $(n-1)$ -dimensional lattice in U , then two classes (T_1, K) and (T_2, K) are twinable.

E. Three- and Two-Dimensional Crystallography

Three-dimensional crystal classes are listed in Appendix B, Table 5.IV. To name these crystal classes, both Schoenflies' notation and the international notation are used.

Fix an \dagger orthogonal \dagger right-handed system (x, y, z) in V . For each crystal system, define three axes as follows. The first axis is the z -axis. The second is the line $x = y = z$ for the cubic system and the x -axis for the others. The third is the y -axis for orthorhombic, trigonal, and hexagonal systems, and the line $x = y, z = 0$ for the others. Then a crystal class is expressed by its generators with respect to these axes. The symbol k ($k = 1, 2, 3, 4, 6$) denotes the rotation by the angle $2\pi/k$ around the axis; \bar{k} is a composition of the rotation k and the central inversion ($v \mapsto -v, v \in V$), and k/m ($k = 2, 4, 6$) means the composition of k and the reflection about the hyperplane perpendicular to the above axis. The symbol m denotes $\bar{2}$. Usually, 1 is omitted. This is the **full international nota-**

tion and a set of generators can be read from it. The **short international notation** is also used.

Schoenflies' notation consists of the letters C_n, D_n, S_n, T, O , and the subscripts h, v, d . The letters denote the group and order: C_n is the \dagger cyclic group of order n ; D_n is the dihedral group of order $2n$; T is the \dagger tetrahedral group isomorphic to the fourth \dagger alternating group; O is the \dagger octahedral group isomorphic to the fourth \dagger symmetric group; and $S_2 = \bar{1}, S_4 = \bar{4}, S_6 = \bar{3}$. The subscripts h, v, d mean that the group is generated by reflections in the hyperplane which is horizontal, vertical, or perpendicular to a diagonal, respectively. The orientation is appropriately given. For example, in C_{nh} and D_{nh} there is an n th order rotation around the z -axis and the horizontal plane is the xy -plane.

The above notations are also applied to noncrystallographic finite subgroups of the 3-dimensional orthogonal group [13].

The symbol for Bravais type is specified by the name of its crystal system and one of the letters P, A, B, C, F, I . Six kinds of **simple Bravais lattices** are given as follows. A simple lattice is generated by $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ and satisfies the following condition. Denote the lengths of vectors by $a = \|\mathbf{a}\|, b = \|\mathbf{b}\|, c = \|\mathbf{c}\|$, and let $\alpha = \angle(\mathbf{b}, \mathbf{c}), \beta = \angle(\mathbf{c}, \mathbf{a}), \gamma = \angle(\mathbf{a}, \mathbf{b})$ be the angles between the indicated pairs of vectors. For **triclinic P**, $a \neq b \neq c \neq a, \alpha \neq \beta \neq \gamma \neq \alpha$; for **monoclinic P**, $a \neq b \neq c \neq a, \alpha = \gamma = 90^\circ \neq \beta$; for **orthorhombic P**, $a \neq b \neq c \neq a, \alpha = \beta = \gamma = 90^\circ$; for **tetragonal P**, $a = b \neq c, \alpha = \beta = \gamma = 90^\circ$; for **hexagonal P** or **trigonal P**, $a = b \neq c, \alpha = \beta = 90^\circ, \gamma = 120^\circ$; for **cubic P**, $a = b = c, \alpha = \beta = \gamma = 90^\circ$. Let T be one of the above lattices generated by $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$. Then the lattices generated by $\{\mathbf{a}, \mathbf{b}, (\mathbf{b} + \mathbf{c})/2\}$, $\{(\mathbf{c} + \mathbf{a})/2, \mathbf{b}, \mathbf{c}\}$, $\{\mathbf{a}, (\mathbf{a} + \mathbf{b})/2, \mathbf{c}\}$, $\{(\mathbf{a} + \mathbf{b} + \mathbf{c})/2, \mathbf{b}, \mathbf{c}\}$, $\{(\mathbf{a} + \mathbf{b})/2, (\mathbf{b} + \mathbf{c})/2, (\mathbf{c} + \mathbf{a})/2\}$ are called **A, B, C, I, F lattices** determined by T . The original lattice T is called the **primitive lattice**. The **trigonal R** lattice is generated by $\{\mathbf{a}', \mathbf{b}', \mathbf{c}'\}$ satisfying $a' = b' = c', \alpha' = \beta' = \gamma' < 120^\circ, \neq 90^\circ$. Set $\mathbf{a} = \mathbf{a}' - \mathbf{b}', \mathbf{b} = \mathbf{b}' - \mathbf{c}', \mathbf{c} = \mathbf{a}' + \mathbf{b}' + \mathbf{c}'$. Then the hexagonal P lattice generated by $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is defined to be the primitive lattice of the trigonal R lattice. The fourteen Bravais lattices are illustrated in Fig. 3.

An arithmetic crystal class (T, K) is denoted by the symbol of the Bravais lattice T and the symbol of K with respect to the action of K on T . Let the primitive lattice of T be generated by $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$. Put the vector \mathbf{a} on the x -axis, and \mathbf{b} on the xy -plane. The action of K is then determined by the international notation for K , except when $K = \bar{4}2m, 32, 3m, \bar{3}m, \bar{6}m2$, and T is the primitive or I lattice. In these cases two kinds of actions interchanging the second and the third symbols are not equivalent. If such is the case, 1 is inserted into the third position for trigonal systems. Let (K) be

the symbol determined as above and (T) be the symbol of T . Then the action (T, K) is expressed by $((T), (K))$. For example, when $K = 32$, there are three arithmetic crystal classes $(P, 321)$, $(P, 312)$ $(R, 32)$.

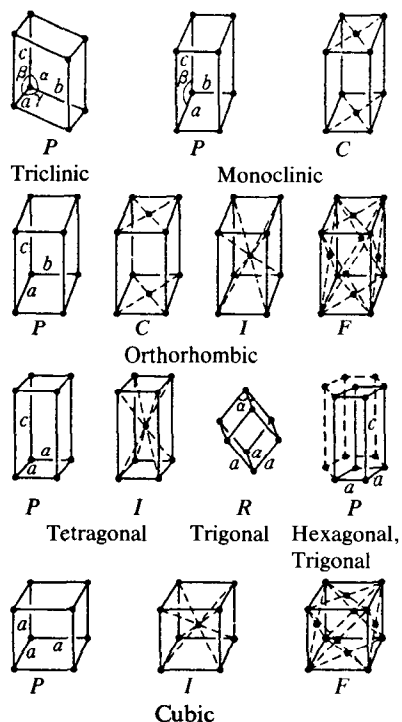


Fig. 3
The conventional unit cell for each of the 14 Bravais lattices.

The notation of space groups determined by (T, K) is given by writing both the symbol (T), and the symbols of actions of elements on V which appear in the symbol (K) defined above. The symbol $(T)(K)$ denotes the 'semidirect product of T and K by the action (T, K) , and is the symmorphic space group determined by (T, K) . Other nonsymmorphic groups are designated by replacing symbols in (K) . If k is replaced by k_j ($j = 1, \dots, k-1$), then k_j is a k -fold screw glide with pitch j/k , that is, the composite of the rotation k and the translation $j\mathbf{f}/k$, where \mathbf{f} is the vector of minimum length of T along the axis of k . If m is replaced by a, b, c, n, d , then it is a **glide reflection**, that is, the composite of the reflection m and the translation in the reflecting hyperplane with the direction of $\mathbf{a}, \mathbf{b}, \mathbf{c}$, a face diagonal and a diamond, respectively. See [6] for the precise meaning of the notation discussed in this paragraph.

The notations for 2-dimensional point groups and space groups can be adopted from the ones for 3-dimensional ones by setting the reference plane as the xy -plane and the z -axis perpendicular to that plane. In this setting, the notation of a plane point group is given by the

corresponding point group in 3-space, and similarly for plane space groups. There is only one **glide operation**, denoted by g . The motion g is the composite of the line reflection about an axis and half the translation along the axis with the vector of minimum length of T . In [13], the precise group-theoretic description of space groups can be found. It should be noted that the notations for the two groups $p3m1$ and $p31m$ are frequently interchanged in the literature (including earlier editions of [13]).

F. History

The first mathematical study of the structure of crystals was done by the mineralogist J. F. C. Hessel (1830). He enumerated the finite subgroups of the 3-dimensional orthogonal group. Afterward, his result was rediscovered many times. A. Bravais (1850) showed that in 3-dimensional Euclidean space there are only 14 different lattices. All space groups consisting of orientation-preserving transformations were determined by C. Jordan (1868, 1869) and L. Sohnke (1879). Finally, almost all the space groups were determined independently by E. S. von Fedorov (1885, 1889) and A. Schoenflies (1887, 1889). Probably as a result of comparison of each other's lists, they established in 1891 the existence of the 230 space groups. Then A. Barlow (1894, 1896) derived these space groups by adding reflection operations to Sohnke's 65 groups.

In 1900 in his eighteenth problem D. Hilbert raised the question of whether the number of equivalence classes of space groups of a given dimension is finite. This was answered affirmatively by L. Bieberbach (1910 [1]). An algorithm for determining space groups was given by H. Zassenhaus (1948 [4]). Their results, including [2, 3], and work by C. Hermann (1949 [9], 1952 [10]) gave a solid foundation for n -dimensional crystallography.

Then followed a concrete treatment of 4-dimensional crystallography. A. C. Hurley (1951), using an earlier work of M. E. Goursat (1889) on finite subgroups of the 4-dimensional special orthogonal group, determined 221 of the 4-dimensional point groups. In the 1960s A. L. Mackay and G. S. Pawley (1963) and others gave 56 of the 4-dimensional Bravais types. After E. C. Dade (1965) determined 9 maximal finite subgroups of $GL(4, \mathbb{Z})$, R. Bülow (1967) and H. Brown, J. Neubüser, and H. Zassenhaus (1968) independently determined 710 arithmetic crystal classes. They reconfirmed the result with H. Wondratschek, and as a result, 64 Bravais types and 227 geometric crystal classes were established (1971 [15]). They ran Brown's computer

program for the determination of space groups based on [4] and obtained 4,783 space groups. Their work culminated in a book (1978 [16]). There have been several attempts to establish a unified treatment of n -dimensional crystallography [5, 8, 14, 17] and crystallography in complex Euclidean spaces.

Two-dimensional space groups were empirically known in ancient times, as demonstrated by artistic decorations. All the 17 2-dimensional space groups appear in the tile patterns of the Alhambra in Granada. A discrete subgroup of the plane motion group whose translation subgroup is only rank one is called a **frieze group**.

The notion of black and white groups was introduced by H. Heesch (1930), H. J. Woods (1935), and A. V. Shubnikov (1951). Only with the introduction of the use of neutron diffraction techniques did it become apparent that these groups could be used in the description of magnetically ordered structures [12]. Color symmetry groups were defined by B. L. van der Waerden and J. J. Bruckhardt (1961 [11]) for an arbitrary number of colors. A mathematical treatment of the twinning structure is found in papers by T. Ito (1938) and R. Sadanaga (1959).

References

- [1] L. Bieberbach, Über die Bewegungsgruppen der Euklidischen Räume, Nachr. Akad. Wiss. Göttingen (1910), 75–84; Math. Ann. 70 (1910), 297–336 and 72 (1912), 400–412.
- [2] L. Bieberbach, Über die Minkowskische Reduktion der positiven quadratischen Formen und die endlichen Gruppen linear ganzzahliger Substitutionen, Nachr. Akad. Wiss. Göttingen (1912), 207–216.
- [3] H. Zassenhaus, Neuer Beweis der Endlichkeit der Klassenzahl bei unimodularer Äquivalenz endlicher ganzzahliger Substitutionsgruppen, Abh. Math. Sem. Univ. Hamburg, 12 (1938), 278–288.
- [4] H. Zassenhaus, Über einen Algorithmus zur Bestimmung der Raumgruppen, Comment. Math. Helv., 21 (1948), 117–141.
- [5] G. Maxwell (and N. Broderick), The crystallography of Coxeter groups, J. Algebra, 35 (1975), 159–177 and 44 (1977), 290–318.
- [6] N. F. M. Henry and K. Lonsdale, International tables for X-ray crystallography, vol. 1, Symmetry groups, Kynoch, Birmingham, 1952, 1965, 1969.
- [7] B. L. van der Waerden, Die Reduktionstheorie der positiven quadratischen Formen, Acta Math., 96 (1956), 265–293.
- [8] R. L. E. Schwarzenberger, Classification of crystal lattices, Proc. Cambridge Philos. Soc., 72 (1972), 325–349.
- [9] C. Hermann, Kristallographie in Räumen beliebiger Dimensionzahl I, Die Symmetrioperationen, Acta Cryst., 2 (1949), 139–145.
- [10] C. Hermann, Translationsgruppen in n -Dimensionen, Zur Struktur und Materie der Festkörper, Springer, 1952.
- [11] B. L. van der Waerden and J. J. Bruckhardt, Farbgruppen, Z. Krist., 115 (1961), 231–234.
- [12] A. V. Shubnikov and N. V. Belov, Colored symmetry, Pergamon, 1964.
- [13] H. S. M. Coxeter and W. O. Moser, Generators and relations for discrete groups, Springer, 1957, 1964, 1972, 1979.
- [14] E. Ascher and A. Janner, Algebraic aspects of crystallography, Helv. Phys. Acta, 38 (1965), 551–571, and Comm. Math. Phys., 11 (1968), 138–167.
- [15] R. Bülow, J. Neubüser, and H. Wondratschek, On crystallography in higher dimensions, Acta Cryst., A27 (1971), 517–535.
- [16] H. Brown, R. Bülow, J. Neubüser, H. Wondratschek, and H. Zassenhaus, Crystallographic groups of 4-dimensional space, Wiley, 1978.
- [17] R. L. E. Schwarzenberger, N -dimensional crystallography, Research notes in math., Pitman, 1980.
- [18] A. Kelly and G. W. Groves, Crystallography and crystal defects, Longman, 1970.
- [19] M. J. Buerger, Introduction to crystal geometry, McGraw-Hill, 1971.
- [20] G. J. Bradley and A. P. Cracknell, The mathematical theory of symmetry in solids, Clarendon Press, 1972.
- [21] C. W. Curtis and I. Reiner, Representation theory of finite groups and associative algebras, Wiley, 1962.

93 (VI.20) Curves

A. Introduction

In the beginning of his *Elements*, Euclid gave definitions such as: A line is a length having no width; an end of a line is a point. However, he left notions such as width and length undefined. Thus his definitions were far from satisfactory. Actually, it was only during the latter half of the 19th century that efforts were made to obtain exact definitions of lines and curves. Euclid, among others, distinguished two kinds of curves: **straight lines** and **curves**. Nowadays, however, **lines** in the sense of Euclid are called curves, and a straight line is considered a curve. A first effort to give an exact definition

of a curve using analytic methods was made by C. Jordan in his *Cours d'analyse I* (1893).

B. Jordan Arcs and Jordan Curves

Following Jordan, we define a **continuous plane curve** C to be the image of a †continuous mapping sending the interval $[0, 1]$ into the Euclidean plane E^2 . Namely, C is the set of points (x, y) in E^2 such that

$$x = f(t), \quad y = g(t), \quad 0 \leq t \leq 1,$$

with continuous functions f, g defined on $[0, 1]$. A continuous curve is also called a **continuous arc**. We call $(f(0), g(0))$ and $(f(1), g(1))$ the **ends** of the arc. Given continuous functions f, g defined on $(0, 1)$, the set $\{(x, y) | x = f(t), y = g(t), 0 < t < 1\}$ is called an **open arc**. More generally, the image of a continuous mapping of $[0, 1], (0, 1), [0, 1)$, or $(0, 1]$ is called an **arc** (or **curve**). Suppose that C is an arc that is the image of an interval I and $P = (x, y)$ is a point on C to which there correspond two elements t_1, t_2 ($t_1 < t_2$) of I such that P is the image of both t_1, t_2 . In this case, the point P is called a **multiple point** on C . An arc having no multiple point is called a **simple arc** or **Jordan arc**.

An arc with one and only one multiple point $P = (f(0), g(0)) = (f(1), g(1))$ is called a **Jordan curve** or **simple closed curve** (\rightarrow Section K). A Jordan curve can be regarded as a topological image in a plane of a circle. Let C be a curve that is the image $\varphi(I)$ of an interval. Then C is said to be of **class C^k (analytic)** if the mapping φ is of †class C^k (†analytic). In general, if S is a topological space, then the image $\varphi(I)$ in S of an interval is called a **curve** in S . In particular, if S has the structure of a differentiable (analytic) manifold, we can define the notion of **curve of class C^k (analytic curve)** in S .

C. Ordinary Curves

A †connected subset of E^2 that is the union of a finite number of simple arcs meeting at a finite number of points is called an **ordinary curve**. An ordinary curve is called a **tree** if it does not contain a subset that is homeomorphic to a Jordan curve. Let p be a point on an ordinary curve C . The †boundary of a sufficiently small †neighborhood of p meets C at a finite number of points, and this number is independent of the choice of the small neighborhood. We call it the **order** of p in C . A point of order 1 is called an **endpoint** of C , a point of order 2 an **ordinary point**, and a point of order ≥ 3 a **branch point**. If we can represent an ordinary curve C as a continuous curve tracing each simple arc of C just once,

we say that C is **unicursal** (Fig. 1). A necessary and sufficient condition for C to be unicursal is that the number of points of odd orders in C be less than or equal to 2 (L. Euler).

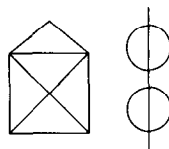


Fig. 1

D. Further Consideration of Definitions

Although the set of ordinary curves as defined above contains most familiar curves, it does not contain a point set defined by $y = \sin 1/x$ in $0 < x \leq 1$ and $-1 \leq x < 0$, and by $-1 \leq y \leq 1$ at $x = 0$ (Fig. 2). This point set is called a **sinusoid**, and it is desirable to obtain a definition of curves wide enough to contain a sinusoid. On the other hand, the notion of continuous curves is, in a sense, too wide, because a curve such as a Peano curve (\rightarrow Section J), which covers a whole square, is among such curves. The notion of simple arcs is too narrow, because even a circle is not a simple arc. As a point set in E^2 , a continuous arc is characterized as a †locally connected †continuum and is sometimes called a **Peano continuum** (H. Hahn, S. Mazurkiewicz). On the other hand, A. Schoenflies, inspired by the statement of the Jordan curve theorem (\rightarrow Section K), considered a closed set that divides the plane into two parts, forming the common boundary of both domains, and called it a closed curve. According to this definition, however, a simple curve is not a closed curve. Thus as a general definition of curves it is not appropriate.

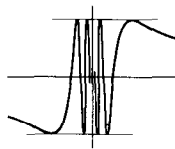


Fig. 2

To give a general notion of curves on a plane (containing sinusoids), we may define a curve as a continuum that is †nowhere dense in E^2 (i.e., a continuum that is a boundary of open sets on the plane) (G. Cantor). Furthermore, to deal with the curves on a topological space, P. S. Uryson and K. Menger defined a **general curve** to be a 1-dimensional continuum (Menger, 1921–1922 [1]). In E^2 , the latter notion coincides with the notion of curves defined by Cantor, while in E^3 a general curve

is a continuum that does not divide any domain (\rightarrow 79 Connectedness).

E. Universal Curve

Consider a 3-dimensional cube I^3 ($I = [0, 1]$). Draw two planes parallel to each face so that the two planes parallel to a face trisect the edges of the cube meeting the planes. Thus I^3 is divided into 3^3 cubes. Let M_1 be the closure of the subset of I^3 that is obtained from I^3 by deleting the cube I_1^3 ($I_1 = [1/3, 2/3]$) and the 6 cubes having common faces with I_1^3 . Then M_1 consists of 20 cubes (Fig. 3). We apply to each of the 20 cubes forming M_1 the same operation that we applied to I^3 and denote by M_2 the union of point sets thus obtained (consisting of 20^2 cubes, the length of whose edges equals $1/3^2$). Repeating this process, we obtain a point set M_n consisting of 20^n cubes, the length of whose edges equals $1/3^n$. Thus we obtain the sequence $M_1 \supset M_2 \supset M_3 \supset \dots$. The set $U = \bigcap_{n=1}^{\infty} M_n$ is a general curve in the sense of Uryson and Menger. Moreover, we can prove that an arbitrary general curve is homeomorphic to a subset of U . Hence we call U the **universal curve**.

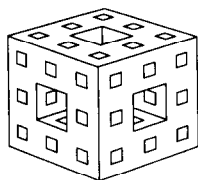


Fig. 3

F. Length of a Curve

In this section, by a curve we mean a continuous curve in a Euclidean space E^n . Let C be a curve in E^n defined by $x_i = f_i(t)$ ($i = 1, 2, \dots, n$; $a \leq t \leq b$); the f_i are real-valued continuous functions defined in $[a, b]$. (We sometimes write this simply as $X = \bar{f}(t)$, where $X = (x_1, \dots, x_n)$.) We divide $[a, b]$ arbitrarily and denote the dividing points by $a = t_0 < t_1 < t_2 < \dots < t_r = b$. Let $X_k = \bar{f}(t_k)$, $k = 0, \dots, r$, and let $\overline{X_{k-1}X_k}$ be the length of the straight line segment joining X_{k-1} and X_k . If the length $l = \sum_{k=1}^r \overline{X_{k-1}X_k}$ of the broken line $(X_0X_1 \dots X_r)$ (Fig. 4) is bounded for any subdivision of $[a, b]$, C is called a **rectifiable curve**, and the upper limit of l with respect to the subdivisions is called the **length** of C . For C to be rectifiable, it is necessary and sufficient that the f_i ($i = 1, 2, \dots, n$) be of bounded variation (Jordan). Thus if C is a rectifiable curve, then each $f_i(t)$ ($i = 1, 2, \dots, n$) is almost everywhere differentiable (H. Lebesgue). In particular, if C

is of class C^1 , then C is rectifiable, and its length can be represented by

$$s = \int_a^b \left(\sum_{i=1}^n \left(\frac{df_i}{dt} \right)^2 \right)^{1/2} dt$$

(\rightarrow 246 Length and Area).

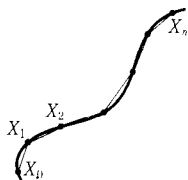


Fig. 4

G. Shapes of Curves

In this section by a curve we mean the image $\varphi(I)$ in a Euclidean space E^n of an interval I (bounded or unbounded), where φ is a continuous mapping. When a curve C of class C^k is given in a Euclidean space E^n , it often becomes necessary to examine the shape of C globally. The determination of the global shape of C from the equation of the curve is called the **curve tracing** of C . The problem has been thoroughly studied in the particular case in which $n = 2$ and the equation of C is given by $f(x, y) = 0$ in a rectangular coordinate system (by $F(r, \theta) = 0$ in a polar coordinate system), where f (or F) is an analytic function. The problems in the case of a rectangular coordinate system are as follows:

Let φ be a single-valued analytic function and I a (bounded or unbounded) interval on the x -axis. If a curve C_0 represented by $y = \varphi(x)$ ($x \in I$) is a subset of C , then C_0 is called a **branch** of C . According as I is bounded or unbounded, C_0 is said to be a **finite branch** or **infinite branch** of C . When a curve represented by $x = \psi(y)$ ($y \in J$) (ψ is a single-valued analytic function and J is an interval on the y -axis) is a subset of C , it is also called a branch of C . If it is necessary to distinguish these two branches, we call the former the **x -branch** and the latter the **y -branch**. C consists of an at most denumerable set of branches. If $P(x_0, y_0) \in C$ and $f_y = \partial f / \partial y \neq 0$ at P , then there exists an x -branch containing P ; if $f_x = \partial f / \partial x \neq 0$, there exists a y -branch of C . If $\partial f / \partial x = 0$, $\partial f / \partial y = 0$ at P , then P is called a **singular point** of C . Points on C that are not singular points are called **ordinary points** of C .

When P is an ordinary point of C , a branch C_0 of C containing P is determined, and the **tangent line** and **normal line** to C at P are the same as those to C_0 and are uniquely determined. The equations of these are $(x - x_0)f_x(x_0, y_0) + (y - y_0)f_y(x_0, y_0) = 0$ and $(x - x_0)f_y(x_0, y_0) - (y - y_0)f_x(x_0, y_0) = 0$, respec-

tively. If we choose a coordinate system with P as origin and the tangent line and normal line as ξ -axis and η -axis, respectively, then the equation of C with respect to this coordinate system is of the form $\eta = c_2\xi^2 + c_3\xi^3 + \dots$ in the neighborhood of P . If we denote by ρ the curvature of C at P , then $\rho = 2c_2 = -(f_{xx}f_y^2 - 2f_{xy}f_xf_y + f_{yy}f_x^2)/(f_x^2 + f_y^2)^{3/2}$. When $c_2 = \rho = 0$, P is a stationary point. A stationary point of a curve of class C^2 on C is also called a **point of inflection**. When P is not a point of inflection and (ξ, η) are points on C in a neighborhood V of P , the sign of η is definite if V is small enough (Fig. 5). However, if P is a point of inflection and $c_3 \neq 0$, then C is of the shape shown in Fig. 6. At a point of inflection, if $c_3 = \dots = c_{v-1} = 0$, $c_v \neq 0$, and v is even, C is of the shape shown in Fig. 5, and if v is odd, C is of the shape shown in Fig. 6.

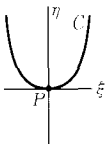


Fig. 5

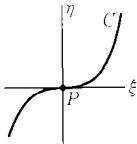


Fig. 6

In a neighborhood of a singular point, C takes various shapes. For example, consider a curve represented by $y^2 = x^2(x + a)$, and let P be the origin $(0, 0)$. If $a > 0$, then there are two branches of C passing through P , and they have different tangents at P (Fig. 7). As in this case, if there are a finite number of different branches passing through P with different tangents, P is called a **node** of C . If $a < 0$, then $P \in C$, but there is no other point of C in the neighborhood of P (Fig. 8). Such a point is called an **isolated point** of C . If $a = 0$, then there are two branches of C starting from P , and the tangents to these at P are the same (Fig. 9). Such a point is called a **cusp** of C . When C is

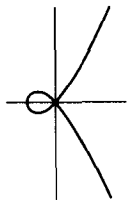


Fig. 7

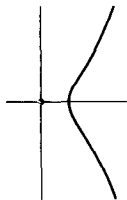


Fig. 8

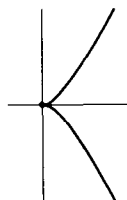


Fig. 9

an algebraic curve, we can examine the shape of a curve in a neighborhood of a singular point using the Puiseux series.

When C has an infinite branch C_0 (for example, when $C_0: y = \varphi(x)$ ($x \in I$) is an x -branch of C and $I = [a, \infty)$), if the tangent to C_0 at $P(x_0, y_0)$ ($x_0 \in I$) has a limiting line for $x_0 \rightarrow \infty$, then the limiting line l is called an **asymptote** of C_0 . In this case, the distance from a point $P(x_0, y_0)$ of C_0 to l converges to zero when $x_0 \rightarrow \infty$. An asymptote of an infinite branch of C is also called an asymptote of C .

H. Special Plane Curves

The following are the well-known curves (for ellipses, parabolas, and hyperbolas → 78 Conic Sections).

Among curves of the third order, those having an equation of the form

$$y^2 = f(x)/(x - a) \tag{1}$$

($f(x)$ is a rational expression of at most the third order in x) are symmetric with respect to the x -axis and have $x = a$ as an asymptote. In particular, if $a > 0$, $f(x) = -x^3$ in (1), then the curve is as shown in Fig. 10 and the origin is a cusp. Let a half-line starting from the origin meet the curve, the circle with diameter $[0, a]$, and the straight line $x = a$ at points X , Y , and A , respectively. Then we have $OX = YA$. This curve is called a **cisoid of Diocles**.

If $a = 0$, $f(x) = c^2(c - x)$ ($c > 0$) in (1), then the curve takes the shape shown in Fig. 11. Let A , C be the points whose coordinates are $(a, 0)$, $(c, 0)$ ($0 < a < c$), respectively, and let X , Y be the points in the first quadrant at which the straight line parallel to the y -axis and passing through A meets the curve and the circle with diameter OC , respectively (Fig. 11). Then we have $AX : AY = OC : OA$. This curve is called a **witch of Agnesi**.

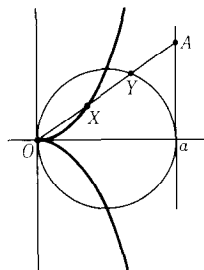


Fig. 10

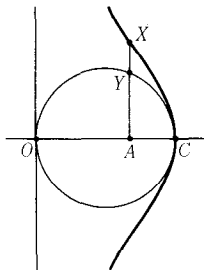


Fig. 11

If $a < 0$, $f(x) = -x^2(x/3 + a)$ in (1), then the curve takes the shape shown in Fig. 12a. If we rotate it by $\pi/4$ and put it in the position shown in Fig. 12b, then the equation of the curve takes the form $x^3 + y^3 = 3cxy$ ($c = -\sqrt{2}$

a). If we take as parameter $t = y/x$, then we get the parametric representation $x = 3ct/(1 + t^3)$, $y = 3ct^2/(1 + t^3)$. This curve is called a **folium cartesii** (or **folium of Descartes**). A curve that has a parametric representation of the form $x = \varphi(t)$, $y = \psi(t)$, where φ, ψ are rational functions, is called an (algebraic) **unicursal curve** (or **rational curve**). Such a curve is an algebraic curve of \dagger genus 0.

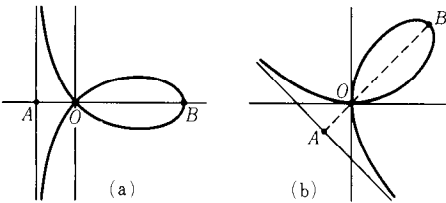


Fig. 12

Let $r = f_1(\theta), r = f_2(\theta), \dots, r = f_k(\theta)$ be equations of curves C_1, C_2, \dots, C_k with respect to a polar coordinate system with origin O . A curve C having equation $r = \lambda_1 f_1(\theta) + \dots + \lambda_k f_k(\theta)$ (the λ_i are constants, usually $+1$ or -1) in the same coordinate system is called a **cissoidal curve** with respect to O . (Fig. 13: $r = -f_1(\theta) + f_2(\theta)$). In Fig. 10, let C_1 be the circumference of the circle with the diameter $[0, a]$, let C_2 be the straight line $x = a$, and put $\lambda_1 = -1, \lambda_2 = 1$. Then we have a cissoid of Diocles. We can regard the folium cartesii as a cissoidal curve obtained from a straight line and an ellipse. When C_1 is a circle with center at O , we call C a **conchoidal curve** of C_2 with respect to O . In particular, when C_2 is a straight line and O is not on C_2 , the conchoidal curve is called a **conchoid of Nicomedes**.

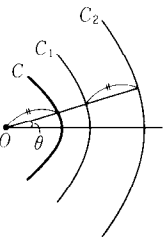


Fig. 13

As shown in Fig. 14, when C_2 is perpendicular to the initial line of the polar coordinate system, the equation of the conchoid is $r = a \sec \theta \pm b$ (b is the radius of C_1), and the Cartesian equation of the curve is $(x - a)^2(x^2 + y^2) = b^2x^2$. According as $a > b, a = b$, or $a < b$, the curve has a node, cusp, or isolated point, respectively. When C_2 is a circle and O is on C_2 , the conchoidal curve of C_2 with respect to O is called a **limaçon** (or **limaçon of Pascal**) (Fig. 15). The equation of a limaçon C with respect to a polar coordinate system having

the diameter of a circle passing through O as its initial line is $r = a \cos \theta \pm b$, while the equation of C with respect to a Cartesian coordinate system is $(x^2 + y^2 - ax)^2 = b^2(x^2 + y^2)$. In this case, if $a > b$, O is a node of the curve; if $a = b$, O is a cusp. When $a = b$, the curve is called a **cardioid** (the curve shown by a dashed line in Fig. 15; see also Fig. 26 below).

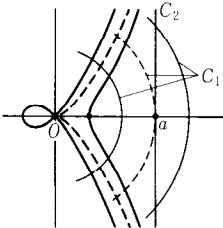


Fig. 14

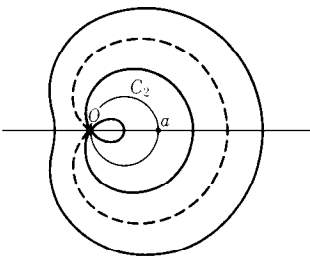


Fig. 15

The locus of a point X having a constant product of its distances from two fixed points A, B is called **Cassini's oval** (Fig. 16). The equation of this curve with respect to the Cartesian coordinate system whose origin O is the midpoint of the segment AB and whose x -axis is the straight line AB is $(x^2 + y^2)^2 - 2a^2(x^2 - y^2) = k^4 - a^4$ (where $AB = 2a, k^2 = AX \cdot BX$). In particular, if $a^2 = k^2$, then O is a nodal point of the curve. In this case, the curve (shown by the dashed line in Fig. 16) is called a **lemniscate** (or **Bernoulli's lemniscate**) (Jakob Bernoulli).

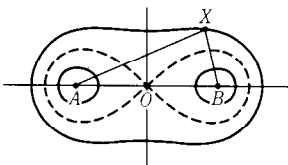


Fig. 16

The locus of the foot of the perpendicular drawn from a fixed point O to the tangent of a fixed curve C at each point of the curve is called the **pedal curve** of C with respect to O . The pedal curve of a rectangular hyperbola with respect to its center is a lemniscate (Fig. 17), and the pedal curve of a circle with respect to a point is a limaçon.

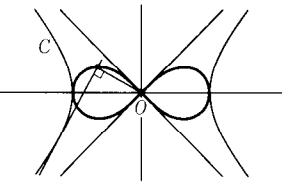


Fig. 17

When a curve C' rolls on a fixed curve C without slipping and is always tangent to C , the locus Γ of a point X kept fixed with respect to the curve C' is called a **roulette** whose **base** is C , **rolling curve** is C' , and **pole** is X . In particular, when C is a straight line, C' is a circle, and X is on C' , Γ is called a **cycloid** (Fig. 18). When X is not on C' , Γ is called a **trochoid** (Fig. 19). A trochoid is represented parametrically by the equations $x = a\theta - b \sin \theta$, $y = a - b \cos \theta$, where the parameter θ is the angle of rotation of C' . When $a = b$, the equation represents a cycloid. The **evolute** and **involute** of a cycloid are also cycloids (Fig. 20).

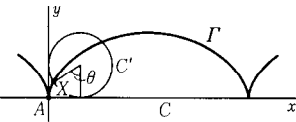


Fig. 18

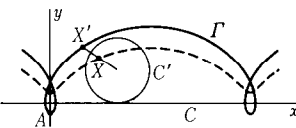


Fig. 19

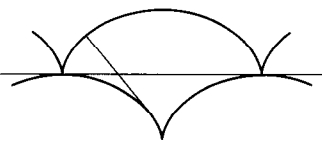


Fig. 20

Suppose that we are in a gravitational field with a given path represented as a cycloid, as is shown in Fig. 21. Assuming that there is no friction, the time necessary for a particle to slide down the path from a point X on the curve to the lowest point C of the curve is independent of the initial position X (C. Huygens). Because of this property, the cycloid is also called a **tautochrone**. Suppose that a particle starts from a point A in the space and slides down to a lower point B along a curve Γ (without friction) under the effect of a gravitational force. To minimize the elapsed time, we simply take Γ as a cycloid that lies in the vertical plane containing AB and has a horizontal line through A as the base (Fig. 22). Because of this property, the cycloid is called

also a **brachistochrone**, i.e., the **line of swiftest descent** (Johann Bernoulli and others).

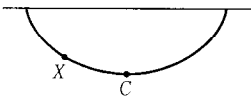


Fig. 21

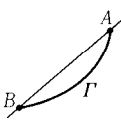


Fig. 22

When the base curve C and the rolling curve C' are both circles and X is on C' , we call Γ an **epicycloid** if C , C' are externally tangent (Fig. 23), and a **hypocycloid** if C , C' are internally tangent (Fig. 24). When X is not on C' , corresponding to these two cases, we have an **epitrochoid** and a **hypotrochoid**, respectively. Let a , b be radii of C , C' , respectively, c the distance from the center of C' to X , and θ the angle of rotation of C' . Then the parametric equations of these curves are $x = (a \pm b) \cos \theta \mp c \cos((a \pm b)/b)\theta$, $y = (a \pm b) \sin \theta - c \sin((a \pm b)/b)\theta$. (Take the upper signs when the curve is an epicycloid and the lower signs when the curve is a hypocycloid. When $b = c$ the equations are equations of epicycloids and hypocycloids.) When the ratio $a : b$ is a rational number p/q (p , q are mutually prime), then C' returns to its initial position after rotating q times around C ; in this case each Γ becomes

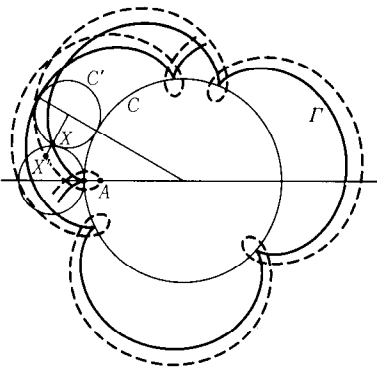


Fig. 23

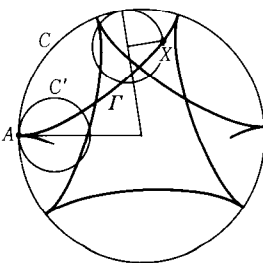


Fig. 24

an algebraic curve. In particular, when $a = 4b = 4c$, the hypocycloid is called an **astroid** (Fig. 25). Its equation (with respect to a Cartesian coordinate system) is $x^{2/3} + y^{2/3} = a^{2/3}$. The envelope of line segments of length a whose endpoints are on the x -axis and y -axis, respectively, is an astroid (Fig. 25). When $a = b = c$ an epicycloid is a **cardioid** (Fig. 26).

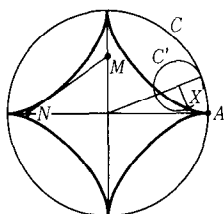


Fig. 25

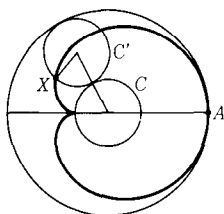


Fig. 26

When the base C is a straight line, the rolling curve C' is an ellipse or a hyperbola, and the pole is a focus of C' , then the roulette is called a **Delaunay curve** (Fig. 27).



Fig. 27

When C is a straight line, C' is a parabola, and X is the focus of C' , the roulette is called a **catenary** (Fig. 28). When we hold two ends of a string of homogeneous density in the gravitational field, the string takes the form of this curve. The equation of the catenary with respect to a Cartesian coordinate system is $y = a \cosh x/a = a(e^{x/a} + e^{-x/a})/2$. The involute starting at the point $A(0, a)$ of this curve is called a **tractrix** (Fig. 29). Let Q be the point of intersection of the tangent at P to the tractrix and the x -axis; then the length of PQ is constant and is equal to a . Consequently, when we drag a weight at A by a string of length OA along the x -axis, the curve described by the weight is the tractrix. The parametric equations of the tractrix are $x = a(\log \tan t/2 + \cos t)$, $y = a \sin t$.

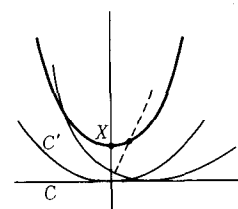


Fig. 28

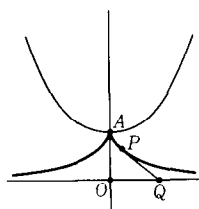


Fig. 29

Suppose that a point Q moves with constant velocity on the x -axis and another point P also moves with constant velocity always toward Q .

The locus of the point P is called a **curve of pursuit** (Fig. 30). When the velocity of Q is α times that of P , the equation of the curve of pursuit is $2(x-a) = y^{1-\alpha}/c(1-\alpha) - cy^{1+\alpha}/(1+\alpha)$ if $\alpha \neq 1$ and $2(x-a) = (1/c)\log y - cy^2/2$ if $\alpha = 1$. We can consider similar problems when Q moves on a general curve instead of on the x -axis.

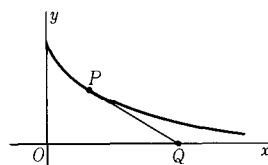


Fig. 30

Many plane curves that are called **spirals** can be expressed by $r = f(\theta)$ (f monotonic) in polar coordinates (r, θ) . An **Archimedes spiral** is a curve having the equation $r = a\theta$ (Fig. 31). Archimedes found that the area bounded by two straight lines $\theta = \theta_1$, $\theta = \theta_2$ ($\theta_1 < \theta_2$) and the curve is $a^2(\theta_2^3 - \theta_1^3)/6$. A **logarithmic spiral** (equiangular spiral or Bernoulli spiral) is a curve having the equation $r = ke^{a\theta}$ (Fig. 32). The angle between the straight line $\theta = \text{constant}$ and the tangent to the curve is constant. Johann Bernoulli found that the involute and evolute of this curve are congruent to the original curve. A curve having the equation $r = a/\theta$ is called a **hyperbolic spiral** (or **reciprocal spiral**), and the one having the equation $r^2\theta = a$ is called a **lituus**. These two spirals are shown in Figs. 33 and 34, respectively. Let $\rho = \varphi(s)$ be the natural equation of a curve. Properties of the curves for which the functions $\varphi(s)$ are simple have been investigated. Specifically, a curve having the natural equation $\rho = ks$ (k is a constant) is a logarithmic spiral. A curve having the equation $\rho = a^2/s$ is called a **Cornu spiral** (or **clothoid**); → 167 Functions of Con-

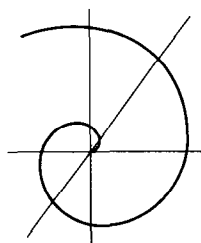


Fig. 31

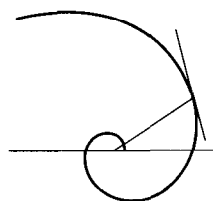


Fig. 32

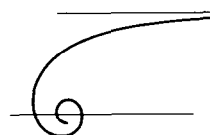


Fig. 33

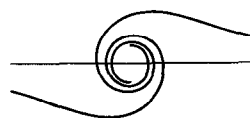


Fig. 34

fluent Type). Its parametric representation is

$$x = a\sqrt{\pi} \int_0^t \cos \frac{\pi t^2}{2} dt,$$

$$y = a\sqrt{\pi} \int_0^t \sin \frac{\pi t^2}{2} dt$$

(†Fresnel integral). M. A. Cornu used this curve in the representation of diffraction in physical optics.

There are also curves that appear as graphs of elementary functions. For example, a curve having the equation $y = \sin x$ is called the **sine curve**, and the graphs of equations $y = e^x$ and $y = \log x$ are called the **exponential curve** and the **logarithmic curve**, respectively, although they are congruent. In contrast to algebraic curves, these analytic curves that are not algebraic are called **transcendental curves**. (Regarding the differential geometric properties of plane and spaces curves → 111 Differential Geometry of Curves and Surfaces; for plane algebraic curves → 9 Algebraic Curves.)

I. Envelopes

Let $\tilde{f}(s, t)$ be a function of class C^1 of real variables s, t . If we fix $t = t_0$, then $r = \tilde{f}(s, t_0)$ is the equation of a curve C_{t_0} with a parameter s . If $s(t)$ is a function of t , then $\tilde{f}(s(t_0), t_0)$ represents a point P_{t_0} on C_{t_0} . Let E be the locus of P_{t_0} when t_0 moves. If $s(t)$ is a function of class C^1 , then E is a curve of class C^1 . If E and C_{t_0} are always tangent at each point P_{t_0} , we call E the **envelope** of the family of curves C_t . When $\tilde{f}(s, t)$ is given, to find E we need only determine the function $s(t)$. We note that $\partial \tilde{f} / \partial s = \lambda \partial \tilde{f} / \partial t$ is a condition that must be satisfied by the function $s(t)$. When $n = 2$ and the equation of C_{t_0} is given in the form $f(x, y, t_0) = 0$, the point of intersection of C_{t_0} and $f_t(x, y, t_0) = 0$ ($f_t = \partial f / \partial t$) is P_{t_0} . The equation $R(x, y) = 0$ obtained by eliminating t from $f(x, y, t) = 0$ and $f_t(x, y, t) = 0$ is called the **discriminant** of $f(x, y, t) = 0$. The set of points (x, y) satisfying the discriminant $R = 0$ is the union of E and the locus of the singular points of C_{t_0} .

J. Peano Curve

A †continuous curve in the Euclidean plane E^2 (i.e., the image $f(I)$ of a segment $I = [0, 1]$ under a continuous mapping $f: I \rightarrow E^2$) may cover a square. We call such a curve a **Peano curve** after G. Peano, who gave the first example. D. Hilbert simplified the example and constructed a Peano curve as follows (*Math. Ann.*, 36 (1890), 38 (1891)).

Divide a square and a segment into four

equal parts (Fig. 35) and let each square D_i correspond to the segment T_i ($i = 0, 1, 2, 3$). Then divide each D_i into four equal parts (Fig. 36) and let D_{ij} correspond to T_{ij} ($j = 0, 1, 2, 3$), and continue this process (Fig. 37). A sequence of squares $D_i \supset D_{ij} \supset D_{ijk} \supset \dots$ has the unique common point $p_{ijk\dots}$, and we let this point correspond to the unique common point $t_{ijk\dots}$ of the sequence of segments $T_i \supset T_{ij} \supset T_{ijk} \supset \dots$. The correspondence $t_{ijk\dots} \rightarrow p_{ijk\dots}$ is a continuous mapping of the segment $[0, 1]$ onto the square D , and this continuous curve has double points, triple points, and quadruple points. The set of multiple points has the cardinal number of the continuum and is a †dense set in the square. We can improve this method so that there are no multiple points other than double and triple points, but it is impossible to eliminate triple points altogether.

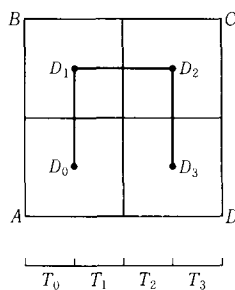


Fig. 35

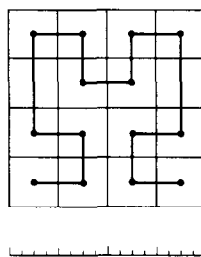


Fig. 36

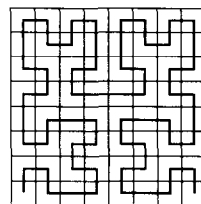


Fig. 37

We can also construct Peano curves as shown in Figs. 38, 39, and 40. That is, we “bisect” a triangle and a segment $[0, 1]$ successively and build a correspondence as follows: Let a point represented by a binary number $t = 0.ijk\dots$ ($i, j, k, \dots = 0, 1$) correspond to the unique common point of the sequence

$\Delta_i \supset \Delta_{ij} \supset \Delta_{ijk} \supset \dots$ and obtain a continuous mapping from the segment $[0, 1]$ onto the triangle.

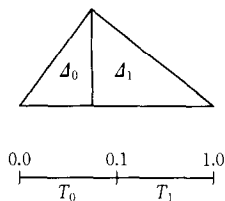


Fig. 38

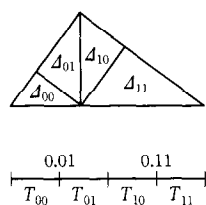


Fig. 39

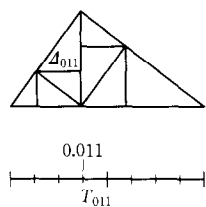


Fig. 40

K. The Jordan Curve Theorem

The **Jordan curve theorem** states: A Jordan curve J in the plane \mathbf{R}^2 separates \mathbf{R}^2 into inner and outer †regions (C. Jordan, *Cours d'analyse*, 2nd ed., 1893). More precisely, $\mathbf{R}^2 - J$ is the disjoint union $G_1 \cup G_2$ of two regions G_1 and G_2 whose common †boundary is J . Let p be a point of J . Then there is a Jordan arc with p as an endpoint such that all points of the Jordan arc are contained in G_i ($i = 1$ or 2) except for p (A. Schönflies), that is, J is **accessible** from G_i . Conversely, let J be a compact subset of \mathbf{R}^2 such that $\mathbf{R}^2 - J = G_1 \cup G_2$ and $G_1 \cap G_2 = \emptyset$, where the G_i are regions such that J is accessible from both G_1 and G_2 . Then J is a Jordan curve (Schönflies, 1908). A homeomorphism between a Jordan curve J and the circle C extends to a homeomorphism (more precisely to a †conformal mapping) between a plane containing J and a plane containing the circle C (\rightarrow 65 Combinatorial Manifolds G; 77 Conformal Mappings).

References

- [1] K. Menger, *Kurventheorie*, Teubner, 1932 (Chelsea, 1967).

- [2] R. L. Wilder, *Topology of manifolds*, Amer. Math. Soc. Colloq. Publ., 1949.
 [3] G. T. Whyburn, *Analytic topology*, Amer. Math. Soc. Colloq. Publ., 1942.
 [4] G. Loria, *Spezielle algebraische und transzendenten ebene Kurven. Theorie und Geschichte I, II*, Teubner, 1910, 1911.
 [5] F. G. Teixeira, *Traité des courbes spéciales remarquables planes et gauches I–III*, Coimbre, 1908–1915.
 [6] P. S. Alexandrov and H. Hopf, *Topologie I*, Springer, 1935.
 [7] S. Lefschetz, *Introduction to topology*, Princeton Univ. Press, 1949.
 [8] M. H. A. Newman, *Elements of the topology of plane sets of points*, Cambridge Univ. Press, second edition, 1951.
 [9] M. Greenberg, *Lectures on algebraic topology*, Benjamin, 1966.
 [10] E. E. Moise, *Geometric topology in dimensions 2 and 3*, Springer, 1977.

94 (X.11) Curvilinear Integrals and Surface Integrals

A. General Remarks

The integral of a function (or more precisely, a †differential form) along a †curve (†surface) is called a **curvilinear integral (surface integral)**. Because a curvilinear integral is a special case of the Stieltjes integral, we shall first explain this notion, formulated by T. J. Stieltjes (1894) as a generalization of the †Riemann integral. The notion was introduced in connection with Stieltjes's study of †continued fractions, and led to the idea of integrals with respect to general measures.

B. The Riemann-Stieltjes Integral

Suppose that $f(x)$, $\alpha(x)$ are real-valued bounded functions defined on $[a, b]$. Take a partition of the interval $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$ (\rightarrow 216 Integral Calculus) and consider the Riemann sum with respect to $\alpha(x)$:

$$\sum_{i=0}^{n-1} f(\xi_i)(\alpha(x_{i+1}) - \alpha(x_i)), \quad \xi_i \in [x_i, x_{i+1}].$$

Suppose that the Riemann sum tends to a fixed number as $\max(x_{i+1} - x_i)$ tends to zero. Then the limit is called the **Riemann-Stieltjes integral** (or simply **Stieltjes integral**) of $f(x)$ with respect to $\alpha(x)$ and is denoted by

$\int_a^b f(x) d\alpha(x)$. The Riemann integral of $f(x)$ is a special case, where $\alpha(x) = x$.

The Riemann-Stieltjes integral has the elementary properties, such as linearity, of the usual Riemann integral. We also have the following theorem: The integral $\int f(x) d\alpha(x)$ exists for every continuous function $f(x)$ if and only if $\alpha(x)$ is of \dagger bounded variation. Hence when we consider the Stieltjes integral of $f(x)$ with respect to $\alpha(x)$, we usually assume that $f(x)$ is continuous and $\alpha(x)$ is of bounded variation. However, the Stieltjes integral can be defined if $f(x)$ is of bounded variation (not necessarily continuous) and $\alpha(x)$ is continuous (not necessarily of bounded variation). If a sequence $f_n(x)$ ($n = 1, 2, \dots$) of uniformly bounded continuous functions defined on the interval $[a, b]$ converges to a continuous function $f(x)$ on the interval $[a, b]$, we have

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) d\alpha(x) = \int_a^b f(x) d\alpha(x),$$

where $\alpha(x)$ is a function of bounded variation. Furthermore, if $\alpha(x)$ and $\alpha_n(x)$ ($n = 1, 2, \dots$) are functions of bounded variation whose \dagger total variations are uniformly bounded and $\lim_{n \rightarrow \infty} \alpha_n(x) = \alpha(x)$ at every point of continuity of $\alpha(x)$, then we have

$$\lim_{n \rightarrow \infty} \int_a^b f(x) d\alpha_n(x) = \int_a^b f(x) d\alpha(x)$$

for every continuous function $f(x)$ on $[a, b]$ (**Helly's theorem**).

Let $\alpha(x)$ be a \dagger strictly monotone increasing continuous function, and let $\beta(y)$ be its inverse function. Then we have

$$\int_a^b f(x) d\alpha(x) = \int_{\alpha(a)}^{\alpha(b)} f(\beta(y)) dy, \quad (1)$$

where the right-hand side is the usual Riemann integral. A function $\alpha(x)$ of bounded variation is represented as the difference of two strictly monotone increasing functions $\alpha_1(x)$ and $\alpha_2(x)$. If we denote by $\beta_i(y)$ the inverse function of $\alpha_i(x)$ ($i = 1, 2$), we have

$$\begin{aligned} \int_a^b f(x) d\alpha(x) &= \int_{\alpha_1(a)}^{\alpha_1(b)} f(\beta_1(y)) dy \\ &\quad - \int_{\alpha_2(a)}^{\alpha_2(b)} f(\beta_2(y)) dy. \end{aligned} \quad (2)$$

If $\alpha'(x)$ exists and is continuous, we have

$$\int_a^b f(x) d\alpha(x) = \int_a^b f(x) \alpha'(x) dx. \quad (3)$$

C. The Lebesgue-Stieltjes Integral

Suppose that $\alpha(x)$ is a monotone increasing and right continuous function and $I = (x_1, x_2]$.

We define an \dagger interval function $U(I) = \alpha(x_2) - \alpha(x_1)$. It is nonnegative and countably additive. Hence by utilizing $U(I)$ we can construct the outer measure and also a completely additive measure (\rightarrow 270 Measure Theory; 380 Set Functions). The Lebesgue integral with respect to this measure is called the **Lebesgue-Stieltjes integral** (or **Lebesgue-Radon integral**) and is denoted by $\int_a^b f(x) d\alpha(x)$. If $\alpha(x)$ is a strictly monotone increasing continuous function and $\beta(y)$ its inverse function, then formula (1) is true if the left-hand side is a Lebesgue-Stieltjes integral and the right-hand side is a Lebesgue integral. If $\alpha(x)$ is a function of bounded variation, decomposing $\alpha(x)$ into the difference of two strictly monotone increasing functions, we also have formula (2). If $\alpha(x)$ is \dagger absolutely continuous, formula (3) is valid, where the right-hand side is a Lebesgue integral.

The Stieltjes integral has the following two properties.

Integration by parts: In the interval $[a, b]$, we have

$$\int_a^b U dV + \int_a^b V dU = U(b)V(b) - U(a)V(a)$$

if one of $U(x)$, $V(x)$ is continuous and the other is of bounded variation.

Second mean value theorem: If $U(x)$ is monotone increasing and $V(x)$ is continuous, then there exists a ξ in $[a, b]$ such that

$$\begin{aligned} \int_a^b U dV \\ = U(a)(V(\xi) - V(a)) + U(b)(V(b) - V(\xi)). \end{aligned}$$

D. The Curvilinear Integral

A continuous mapping from an interval $a \leq t \leq b$ in \mathbf{R}^1 into \mathbf{R}^n : $\varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))$ is an oriented curve. Suppose that a function $f(x_1, \dots, x_n)$ is defined in a neighborhood U of the image C of the mapping $\varphi(t)$ or merely on C . The Stieltjes integral

$$\int_a^b f(\varphi_1(t), \dots, \varphi_n(t)) d\varphi_i(t), \quad i = 1, \dots, n, \quad (4)$$

is called the **curvilinear integral** of the function $f(x_1, \dots, x_n)$ along the curve C with respect to x_i and is denoted by $\int_C f dx_i$. The curve C is called the **contour** (or **path**) of the integration, $\varphi(a)$ is called the **initial point** (or **lower end**), and $\varphi(b)$ is called the **terminal point** (or **upper end**) of the integration. Let C be a rectifiable curve, and denote by $s(t)$ the arc length of C from the initial point to the point $\varphi(t)$. Then the Stieltjes integral

$$\int_a^b f(\varphi(t)) ds(t)$$

or simply $\int_C f ds$ is called the **curvilinear integral with respect to the line element**. Here, the line element, denoted by ds , means $\sqrt{(\varphi'_1(t))^2 + \dots + (\varphi'_n(t))^2} dt$ when $\varphi(t)$ is of class C^1 . If the integrand in (4) is of bounded variation as a function of t , the curvilinear integral is well defined. If C is a rectifiable curve, the curvilinear integral is defined for an arbitrary continuous function. In the usual case, we are concerned mainly with this sort of situation. For a differential form $\omega = f_1 dx_1 + \dots + f_n dx_n$ defined on U , the curvilinear integral $\int_C \omega$ is defined by $\sum_{i=1}^n \int_C f_i dx_i$.

The curvilinear integral is linear with respect to its integrand. If the terminal point of C_1 is the initial point of C_2 , we can construct the joint curve $C = C_1 + C_2$, and we have **additivity for the contours**

$$\int_C f dx_i = \int_{C_1} f dx_i + \int_{C_2} f dx_i$$

(a similar formula holds if we replace dx_i by ds). Monotonicity, which asserts that $\int_C f dx_i \leq \int_C g dx_i$ whenever $f \leq g$, holds if $\varphi_i(t)$ is monotone increasing, and monotonicity also holds for the curvilinear integral with respect to the line element.

If $n=2$, \mathbf{R}^2 can be identified with the complex plane $\mathbf{C} = \{z = x + iy\}$, and we define $\int_C f(z) dz$ by

$$\left\{ \int_C u(x) dx - \int_C v(z) dy \right\} + i \left\{ \int_C v(z) dx + \int_C u(z) dy \right\},$$

where $f(z) = u(z) + iv(z)$. The integral is then said to be an integral in the complex domain. (For the application of integrals in the complex domain to complex analysis → 198 Holomorphic Functions.)

E. The Surface Integral

By an m -dimensional smooth surface S we mean the image S of a regular mapping of class C^1 from a domain G in \mathbf{R}^m into \mathbf{R}^n ($m < n$), $\xi(u) = (\xi_1(u_1, \dots, u_m), \dots, \xi_n(u_1, \dots, u_m))$. Given a continuous function $f(x_1, \dots, x_n)$ defined in a neighborhood U of S in \mathbf{R}^n , the multiple integral

$$\int \dots \int_G f(\xi_1(u), \dots, \xi_n(u)) \times \frac{D(x_1, \dots, x_{i_m})}{D(u_1, \dots, u_m)} du_1 \dots du_m, \quad (5)$$

$\{i_1, \dots, i_m\} \subset \{1, \dots, n\}$,

is called the **surface integral** of f along S with respect to x_{i_1}, \dots, x_{i_m} and is denoted by

$\int_S f dx_{i_1} \dots dx_{i_m}$ or $\int \dots \int_S f dx_{i_1} \dots dx_{i_m}$. This definition depends on the choice of the parameters (u_1, \dots, u_m) in the following way. Let (u'_1, \dots, u'_m) be another parametrization of S . Since

$$\frac{D(x_{i_1}, \dots, x_{i_m})}{D(u_1, \dots, u_m)} = \frac{D(x_{i_1}, \dots, x_{i_m})}{D(u'_1, \dots, u'_m)} \frac{D(u'_1, \dots, u'_m)}{D(u_1, \dots, u_m)}$$

and

$$du'_1 \dots du'_m = \left| \frac{D(u'_1, \dots, u'_m)}{D(u_1, \dots, u_m)} \right| du_1 \dots du_m,$$

thus according as

$$\frac{D(u_1, \dots, u_m)}{D(u'_1, \dots, u'_m)} \geq 0,$$

the foregoing expression remains the same or changes its sign. Usually we assign to the parameters (u_1, \dots, u_m) positive or negative orientation, in which case S is called oriented. If we replace the Jacobian $D(x_{i_1}, \dots, x_{i_m})/D(u_1, \dots, u_m)$ in (5) by the quantity

$$\left(\sum_{i_1 < \dots < i_m} \left(\frac{D(x_{i_1}, \dots, x_{i_m})}{D(u_1, \dots, u_m)} \right)^2 \right)^{1/2},$$

which corresponds to the surface element of S , the integral is called a **surface integral with respect to the surface element** and is denoted by $\int_S f dS$ or $\int_S f d\sigma$. The surface integral of a differential form of degree m in \mathbf{R}^n is similarly defined. In the case $m=1$, the surface integral reduces to a curvilinear integral. Just as the Stieltjes integral is a generalization of the curvilinear integral, there are several ways to generalize the notion of surface integral without assuming that the mapping $\xi(u)$ is of class C^1 .

F. The Stokes Formula

Let S be an m -dimensional smooth surface in \mathbf{R}^n ($m \leq n$) and ∂S be the $(m-1)$ -dimensional surface corresponding to the boundary of S . Let ω be a differential form of class C^1 of degree $(m-1)$ and $d\omega$ be its exterior derivative. Then we have $\int_{\partial S} \omega = \int_S d\omega$, which is called the **Stokes formula** (or the **Green-Stokes formula**). (For the Stokes formula on a general differentiable manifold → 105 Differentiable Manifolds.) As special cases of the Stokes formula, we have the following three classical theorems:

(1) The case of a plane domain: Let D be a bounded domain on the xy -plane bounded by a finite number of smooth curves C with positive directions. If $\omega = P dx + Q dy$ is a differential form of class C^1 on \bar{D} , we have

$$\int_C P dx + Q dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy, \quad (6)$$

Curvilinear Integrals and Surface Integrals

since $d\omega = (\partial Q/\partial x - \partial P/\partial y)dx \wedge dy$. This is called **Green's formula** (or **Green's formula on the plane**). Equality (6) is true if P, Q are totally differentiable and the integrand on the right-hand side is continuous (even if the functions $\partial Q/\partial x, \partial P/\partial y$ themselves are not continuous) (E. Goursat). This formula remains true under the following weaker assumptions: (i) C is rectifiable; (ii) P and Q are continuous in \bar{D} and $\partial P/\partial y$ and $\partial Q/\partial x$ are continuous and summable (in the sense of Lebesgue) in D .

(2) The case of a domain in a 3-dimensional space: Let D be a bounded domain in xyz -space surrounded by a finite number of smooth surfaces S . For a vector field $\mathbf{V} = \langle P, Q, R \rangle$ of class C^1 on \bar{D} , we put $\omega = P dy \wedge dz + Q dz \wedge dx + R dx \wedge dy$. Then since $d\omega = (\partial P/\partial x + \partial Q/\partial y + \partial R/\partial z)dx \wedge dy \wedge dz$, we have

$$\begin{aligned} & \iint_S P dy dz + Q dz dx + R dx dy \\ &= \iiint_D \operatorname{div} \mathbf{V} dx dy dz \\ &= \iiint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx dy dz. \end{aligned} \quad (7)$$

Equality (7) is called the **Gauss formula** (**Ostrogradskii's formula** or the **divergence theorem**). The left-hand side of (7) is equal to the surface integral $\iint_S (\mathbf{V}, \mathbf{n}) d\sigma$, which is the vector flux through S (where \mathbf{n} means the outer unit normal vector of the surface S). This formula remains true under the following weaker assumptions: (i) S is piecewise of class C^1 ; (ii) P, Q, R are continuous in \bar{D} and $\partial P/\partial x, \partial Q/\partial y$, and $\partial R/\partial z$ are continuous and summable in D .

(3) The case of a bordered surface in a 3-dimensional space: Let $\bar{S}: x = x(u, v), y = y(u, v), z = z(u, v)$ ($(u, v) \in \bar{G}$) be a smooth surface in xyz -space, and suppose that the boundary Γ of the domain of the parameters \bar{G} consists of a finite number of smooth curves with positive direction. The boundary C of the surface S is the image curve of Γ . Now let $\mathbf{V} = \langle P, Q, R \rangle$ be a vector field of class C^1 on S , \mathbf{n} be the unit normal vector of S (its direction being canonically assigned by the parameter (u, v)), and \mathbf{t} be the unit tangent vector, and set $\omega = P dx + Q dy + R dz$. Then since $d\omega = (\partial R/\partial y - \partial Q/\partial z)dy \wedge dz + (\partial P/\partial z - \partial R/\partial x)dz \wedge dx + (\partial Q/\partial x - \partial P/\partial y)dx \wedge dy$, we have

$$\begin{aligned} & \iint_S (\operatorname{rot} \mathbf{V}, \mathbf{n}) d\sigma = \iint_S \left(\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dy dz \right. \\ & \quad \left. + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dz dx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy \right) \\ &= \int_C (P dx + Q dy + R dz) = \int_C (\mathbf{V}, \mathbf{t}) ds. \end{aligned} \quad (8)$$

Equation (8) is called the **Stokes formula** (—Appendix A, Table 3.III).

References

- [1] H. Lebesgue, *Leçons sur l'intégration et la recherche des fonctions primitives*, Gauthier-Villars, second edition, 1928, ch. II.
- [2] S. Saks, *Theory of the integral*, Warsaw, 1937.
- [3] D. V. Widder, *The Laplace transform*, Princeton Univ. Press, 1941, ch. I.
- [4] H. Flanders, *Differential forms, with applications to the physical sciences*, Academic Press, 1963.
- [5] H. K. Nickerson, D. C. Spencer, and N. E. Steenrod, *Advanced calculus*, Van Nostrand, 1959.
- [6] H. Okamura, On the surface integral and Gauss-Green's theorem, *Mem. Coll. Sci. Univ. Kyôto*, 26 (1950), 5–14.
- [7] L. Schwartz, *Cours d'analyse*, Hermann, 1967.
- [8] H. Cartan, *Formes différentielles*, Hermann, 1967.

95 (XVI.4)
Cybernetics

The term **cybernetics** was invented in 1947 by Norbert Wiener [1] to denote a field of science that treats the system of control and communication in animals and machine. The term was derived from a Greek word κυβερνήτης, also the source of the word "governor." The stimulus for establishing such a new area of science came from studies of automatic computation, automatic control, and information processing. These fields had given rise to technological innovations such as the high-speed electronic computer and automatic control instruments. Such machines have had a profound influence on the information sciences, as well as on theoretical investigations in biology. However, even now, we can hardly say that cybernetics has been firmly established as a systematic branch of science or of applied mathematics. Nevertheless, it has had far-reaching influence on both biology and machine engineering as a methodology and as a philosophy. In the Soviet Union and in European countries, the word corresponding to cybernetics is still widely used for the "Grenzgebiete" between biology and machine engineering in the wider sense. Nonetheless, it is difficult to say that cybernetics as a whole has undergone systematic development. Systematic theories have

been established and developed for separate parts of the field, and these have become independent disciplines; they are, for example, *control theory, †information theory, the theory of artificial intelligence, the theory of †automata, mathematical biology, genetics, and ecology.

The contributions of Wiener to cybernetics include not only the invention of the concept itself, but also prediction theory or the theory of the Wiener filter [2]. The latter was subsequently formalized by R. E. Kalman from a different point of view as an estimation problem within a linear system, working from finite observation data, and including the non-stationary case (\rightarrow 405 Stochastic Control and Stochastic Filtering G). Another contribution by Wiener is the input-output identification of a nonlinear system using statistical time series analysis of the outputs when the inputs are white noises [4]. This is achieved by expansion of the output function in terms of the convolution of input functions, which corresponds to the expansion of functions in terms of †Hermite interpolation polynomials. The kernel of the expansion is called a **Wiener kernel**. Recently, this method has been applied to input-output identification for nervous systems; it also finds application in nonlinear system theory [5]. Wiener also studied brain waves [6, 9].

Cybernetics in the wider sense, although not necessarily called by this name, may include investigations of the following "Grenzgebiete": information processing of the nervous system [7]; self-organizing systems in the non-equilibrium thermodynamics of Prigogine and Haken; the self-reproducing machines of von Neumann; and the theory of pattern formation in biological or chemical systems [8].

For related topics \rightarrow 176 Gaussian Process I, 395 Stationary Processes D.

References

- [1] N. Wiener, *Cybernetics; or, control and communication in the animal and the machine*, Wiley, 1947; second edition, 1961.
- [2] N. Wiener, *Extrapolation, interpolation and smoothing of stationary time series, with engineering applications*, Wiley, 1949.
- [3] R. E. Kalman and R. S. Bucy, New results in linear filtering and prediction theory, *Trans. ASME, J. Basic Eng.*, 83 (1961), 95–107.
- [4] N. Wiener, *Nonlinear problems in random theory*, MIT Press, 1958.
- [5] Y. W. Lee and S. Schetzen, Measurement of the Wiener kernels of a non-linear system by cross correlation, *Int. J. Control*, 2 (1965), 237–254.

- [6] N. Wiener and A. Rosenblueth, The mathematical formulation of the problem of conduction of impulses in a network of connected excitable elements, specifically in cardiac muscle, *Arch. Inst. Cardiology, Mexico*, 16 (1946), 205–265.
- [7] M. A. Arbib, *Metaphorical brain*, Interscience, 1972.
- [8] P. C. Fife, *Mathematical aspects of reacting and diffusing systems*, Lecture notes in math. 28, Springer, 1979.
- [9] N. Wiener and J. P. Schadé, *Cybernetics of the nervous system*, Progress in brain research XVII, Elsevier, 1965.
- [10] K. Steinbuch and S. W. Wagner, *Neuere Ergebnisse der Kybernetik*, R. Oldenbourg, 1964.
- [11] S. Beer, *Decision and control*, Wiley, 1966.

D

96 (XVI.7) Data Processing

A. General Remarks

With the development of electronic computers, effective systems for data transmission and processing have been created on a large scale, and there is now a large literature concerning data-processing methods and techniques for such systems.

Research on data processing encompasses both specific techniques and entire systems of processing, for example, the system of programming techniques, as well as mathematical problems originating from them, such as the complexity of computations (\rightarrow 71 Complexity of Computations). As domains of application, we have, for example, information retrieval, stock management, program evaluation, and review techniques. We denote the set of data by D and assume that D is finite. Depending on the properties of D , we have various suitable representations and processing problems.

B. The Notion of Data

In recent applications of computers, the main task has been arranging and searching for items or attributes in storage, rather than numerical computation. Individual information is called a **record**. Records arranged and stored in the memories of computers are usually called **data**. A collection of such data is called a **file**. A complex, large-scale data collection is often called a **data base**.

In an abstract sense, the record consists of a string of letters, but it is usually convenient to view it as consisting of its identifying mark followed by a finite sequence of items. The contents can be classified according to their properties, such as topological relations, order relations, or items representing numerical values. For each case, there may be different suitable representations and operations.

We often store information after a suitable process of **information compression**. In some cases the operations are reversible, and complete recovery is possible. In other cases the operations are not completely reversible, and we must throw away part of the information in order to compress the rest. An example of recoverable information is the replacement of a run of 1's (or 0's) in a binary code by its length. An example of lost information often occurs in the graphical manipulation of numbers by means of **hashing**. In this process we

select a suitable function $s = f(a_1, \dots, a_n)$, where the vector (a_1, \dots, a_n) corresponds to the representation of the record in such a way that the value s can be monomorphic for the universe of the records as often as possible, and the data (a_1, \dots, a_n) is kept in the storage corresponding to the value s . Generally, the variable representing the location of data is called a **pointer**.

In the abstract sense, data bases are simply sets of items, but in most cases a base will also exhibit some mathematical structure, such as order relations. In order to handle a data base efficiently, we must take its particular structure into account, and suitably represent that structure. Such structures and their representations are called **data structures**.

C. Linear Structures

The component of a vector or the stations along a single railway line form linearly ordered sets. Their most essential feature is the notion of immediate predecessor or immediate successor. Such a data structure is called a **linear structure**, and the aligned sequence is called an **array**. Multidimensional arrays, such as the elements of a matrix, are stored in the form of 1-dimensional arrays in the memory of a computer.

When the set of data D is a linearly ordered set, it is usually represented as a suitable array. Here, the problems of **ordering** and **table look-up** are fundamental. Ordering in this sense means putting the given elements of D into the order defined for D . Since, historically, sorting machines were used to put punched cards in order, this process is also called **sorting**. The process of arranging several individually sorted data packs into one sorted pack is called **merging**.

A fundamental data-sorting operation is "comparison" with respect to the order for D . There have been many investigations of the estimation of the lower or upper bounds for the number of comparisons and of developing efficient algorithms. Asymptotically, $O(n \log n)$ is the theoretical lower bound for n elements, and some algorithms are known to achieve this bound (\rightarrow 71 Complexity of Computations).

If there is given a univalent correspondence $f: D \rightarrow D'$ and the correspondence table is stored in the memory in a suitable form, the problem of table look-up arises, which requires finding $f(d)$ for a given $d \in D$. In this process, a fundamental operation is a comparison of $d \in D$ with some $x \in D$ in the table. There have been many investigations of efficient arrange-

ments and of algorithms with minimal numbers of comparisons.

D. List Representation

If the given finite set D has an order satisfying the reflexive and transitive laws, addressing is often used as a medium for representing the order in the memory. Let d_1, \dots, d_n be all the immediate successors to an element $d \in D$. We can represent them by the following sequence of triplets:

If $n=0$, we take $(d, 0, 0)$;
 If $n=1$, we take $(d, d_1^*, 0)$;
 If $n=2$, we take (d, d_1^*, d_2^*) ;
 If $n=3$, we take $(d, d_1^*, e_1^*), (e_1, d_2^*, d_3^*)$;
 If $n \geq 4$, we take $(d, d_1^*, e_1^*), (e_1, d_2^*, e_2^*),$
 $(e_2, d_3^*, e_3^*), \dots, (e_{n-2}, d_{n-1}^*, d_n^*),$

where d_i^* means the pointer to d_i , e_i is an element introduced for convenience, and e_i^* is the pointer to e_i . This is called a **list representation**. If we denote by $d \rightarrow d_1$, the fact that d_1 is an immediate successor of d , then the total set can be regarded as a \dagger direct graph. Usually, this graph is a \dagger tree, and then the graph is called a **tree structure** and its representation a **tree representation**. The advantage of this representation is that addition or deletion is quite easy.

The set of logical formulas is partially ordered, with the order given by the rules of inference. The set consisting of a series of inferences forms an ordered subset. Thus, if a tree representation can be automatically treated, so can the process of inference.

In dealing with linguistic data (words or sentences, for example), it is often natural to consider a noncommutative \dagger free semigroup D generated by a finite number of generators (the alphabet or vocabulary). In this case, if there exists a natural order for the generators, it determines in D a lexicographic partial ordering. Then the tree representation can be used for representing a dictionary whose entires are elements of D . This method is not efficient with respect to speed of table look-up and economy of memory, but it sometimes has the advantage of simplifying the treatment of complicated data.

E. Memory Devices for Processing

In dealing with algebraic formulas or languages with parentheses, data maintenance methods such as tree representation or push-down storage are often convenient as auxiliary memory-controlling methods. The character-

istic of the **push-down storage** method is that it returns the data in reverse order with respect to the time of acceptance and remittance.

It is sometimes called a **stack** or a **first-in-last-out memory**. Contrary to a stack, there is a storage that returns the data in the same order as the acceptance. It is called a **queue** or a **first-in-first-out memory**.

F. Information Retrieval

A request to obtain all the records qualified by some property in a data base Ω is called a **query**. The part of the collection of records Ω qualified for a query, therefore, can be identified by its characteristic subset of the space of all possible records. A procedure that determines the subset Ω_q of Ω composed of all the records relevant to a query q is called the **information retrieval** of the query. A system which provides such a procedure for every query $q \in Q$ is called an **information retrieval system** organized for Ω with respect to Q .

It is essential to design the system in such a way that records can be retrieved quickly for queries in a certain class. Such a selected set of queries can be composed of all the first-order queries specifying an item and asking for all the records containing the attribute value or the key characterizing the item. It can, in some cases, include second- or higher-order queries specifying a number of items and asking for all the records in which attribute values or keys characterizing the item occur simultaneously.

In an information retrieval system, the master file of a data base Ω is usually organized by way of auxiliary memories on data structures such as sequential files, indexed sequential files, virtual storage files, or direct access files, using the magnitude of the accession number or primary key of each record. Various management systems, called SAM (Sequential Access Method), ISAM (Index Sequential Access Method), VSAM (Virtual Storage Access Method), or DAM (Direct Access Method), exist.

In addition to the organization of such a master file, various directory files or indexes are organized in order to retrieve each query $q \in Q$ quickly, because there are many relevant keys or combination of keys other than the primary key. An **inverted filing scheme** (IFS) is a typical scheme for organizing such indexes or directory files. A bucket B_i or an addressable set of secondary memories is provided for each canonical query $q_i \in Q$ in a one-to-one way. An index or a list of accession numbers of pertinent records is organized in each of the buckets contiguously so as to make it possible

to retrieve all accession numbers of relevant records quickly. The essentials of an IFS is to define a one-to-one QAT (Query to Address Transformation) from the inverted set of queries $Q = \{q_i\}$ to the set of buckets $B = \{B_i\}$.

An IFS has efficient retrieval performance with respect to the inverted canonical queries. The scheme, however, may in some cases require a large number of buckets. Moreover, the scheme requires quite redundant storage of accession numbers by a number of buckets because a record may be pertinent to a number of canonical queries simultaneously. Although higher retrieval performance can be expected by including higher-order queries in the inverted set, the space and machine time needed for such organization is prohibitive. This is one of the reasons why an IFS for first-order queries is preferable in almost all practical cases. Another reason is that if the first-order queries are inverted, every retrieval can logically be performed by certain Boolean operations executed among a certain number of retrieved sets of accession numbers. A trade-off of space and time needed for the organization of a scheme and its retrieval performance might be the determining factor for the selection of the set of canonical queries to be inverted.

An attempt to overcome the limitations inherent in the inverted scheme can be found in the work on the balanced file organization scheme due to Abraham et al. [7]. By extracting the essentials, Yamamoto et al. [9, 10] defined a BFS (Balanced File-organization Scheme) in a wider sense as follows:

- (i) Buckets are organized in such a manner that every bucket is associated with more than one query.
- (ii) Every canonical query is associated with a unique bucket.
- (iii) The accession number of a record with some additional information is stored in a bucket once if and only if it is pertinent to at least one of the associated queries.

The essentials of a BFS is to define a many-to-one transformation from the set of canonical queries Q to the set of bucket addresses B , or an MQAT (Multiple Queries to Address Transformation). An MQAT defines a partition of Q into mutually disjoint subsets. It is a generalization of a QAT which defines an IFS. If c is the number of queries to be associated simultaneously with a unique bucket in a BFS, then the number of buckets to be prepared is $1/c$, which is a drastic reduction from that for an IFS. This reduction makes it possible to extend the feasible range of canonical queries. Actually, in the system HUNDRED (Hiroshima University New Documents REtrieval and Dissemination), a

51-to-1 MQAT is used for the organization of indexes.

Although some space overload may occur by storing some additional information in order to tell which record is pertinent to which query in a bucket, reduction of redundancy can be expected, because a record may be pertinent to more than one query simultaneously. The reduction in the number of buckets and redundancy would contribute much in saving time and space needed for the organization of indexes.

Among those BFS's defined by MQAT's with a given data base Ω , Q , and the number c of queries associated with the same bucket, one BFS can be called the best if its redundancy is the least under some reasonable assumptions imposed on the distribution of keys. Several combinatorial problems have been raised and solved in this connection [10].

References

- [1] A. T. Berztiss, Data structure—Theory and practice, Academic Press, 1971.
- [2] H. S. Stone, Introduction to computer organization and data structures, McGraw-Hill, 1972.
- [3] A. K. Scidmore and B. L. Weinberg, Storage and search properties of a tree-organized memory system, *Comm. ACM*, 6 (1963), 28–31.
- [4] R. D. Luce, R. R. Bush, and E. Galanter, Handbook of mathematical psychology II, Wiley, 1963.
- [5] Y. Bar-Hillel, Language and information, Addison-Wesley, 1964.
- [6] J. A. Fodor and J. J. Katz, The structure of language: Readings in the philosophy of language, Prentice-Hall, 1964.
- [7] C. T. Abraham, S. P. Ghosh, and D. K. Ray-Chaudhuri, File organization schemes based on finite geometries, *Information and Control*, 12 (1968), 143–163.
- [8] S. P. Ghosh, Data base organization for data management, Academic Press, 1977.
- [9] S. Yamamoto, H. Ikeda, S. Shige-eda, K. Ushio, and N. Hamada, Design of a new balanced file organization scheme with the least redundance, *Information and Control*, 21 (1975), 72–91.
- [10] S. Yamamoto, Design of optimal balanced filing schemes, *Proc. IBM Working Conference on Database Engineering*, in *Database Engineering*, Springer 1981, vol. II, 23–35.
- [11] D. E. Knuth, The art of computer programming, Addison-Wesley, vol. 1, revised edition, 1976; vol. 2, revised edition, 1980; vol. 3, 1973.

97 (I.10) Decision Problem

Suppose that we are given a set S and a proposition $P(x_1, x_2, \dots, x_n)$ for elements x_i of S . Then we have the **problem of universal validity of P** , which is the problem of finding a general **algorithm** (i.e., a finitary procedure) by which we can discern whether $P(x_1, \dots, x_n)$ is true for all n -tuples (x_1, \dots, x_n) . The problem of finding an algorithm by which we can discern the validity of $P(x_1, \dots, x_n)$ for some specifically chosen n -tuples (x_1, \dots, x_n) is called the **problem of satisfiability of P** . These two problems are customarily called **decision problems**. The problems are such that affirmative solution of one of them implies negative solution of the other.

To give a precise definition of decision problems, let us note that a \dagger free semigroup with countable generators can be identified with a subset of the set \mathbb{N} of natural numbers (by virtue of \dagger Gödel numbering; \rightarrow 185 Gödel numbers). On the other hand, if \mathfrak{S} is a given \dagger formal system with countably many symbols, the set of all \dagger formulas in \mathfrak{S} is a subset of the free semigroup generated by the symbols in \mathfrak{S} . Thus the set of all formulas in \mathfrak{S} is identified with a subset of \mathbb{N} . A subset M of \mathbb{N} (or $\mathbb{N} \times \mathbb{N} \times \dots \times \mathbb{N}$) is (**general**) **recursive** if its \dagger representing function is general recursive (\rightarrow 356 Recursive Functions). By using the concept of recursive function, a precise definition of the decision problem can be given as follows: **The decision problem of M is solved affirmatively** if and only if we can obtain effectively a procedure defining the representing function of M , and the function is general recursive. **The decision problem of M is solved negatively** if and only if we can obtain a proof that M is not recursive.

For a set A of formulas in \mathfrak{S} , we let $g(A)$ be the set of all Gödel numbers corresponding to the elements of A . Let A' be the set of formulas in A that are deducible in \mathfrak{S} , and let $\tau(A) = g(A')$. The decision problem of the set A of formulas is said to be solved affirmatively (negatively) if the decision problem of $\tau(A)$ is solved affirmatively (negatively). By refining this concept we arrive at the notion of the **degree of (recursive) unsolvability**. Let A and B be subsets of \mathbb{N} . The relation " A is recursive in B and B is recursive in A " (\rightarrow 356 Recursive Functions) is reflexive, symmetric, and transitive. Hence this relation decomposes the class of all subsets in \mathbb{N} into disjoint nonempty equivalence classes. A and B are defined to have the same degree of unsolvability if they belong to the same equivalence class. Thus the degrees of unsolvability can be identified with

the equivalence classes. The degree of recursive sets is 0 .

The relation $\mathbf{a} \leq \mathbf{b}$ is defined between the degrees \mathbf{a} of A and \mathbf{b} of B to mean " A is recursive in B ." Clearly, for any degree \mathbf{a} , we have $0 \leq \mathbf{a}$. The partially ordered system of degrees constitutes an \dagger upper semilattice.

Research on the decision problem has been done mostly in areas related to the \dagger first-order predicate calculus L^1 and the formal systems on it. We now list some important results.

(I) Results concerning L^1 . The decision problem has been solved negatively for the sets of formulas of the following forms. (Here it is assumed that no function symbols appear and that \mathfrak{A} represents a formula involving no occurrence of \forall , \exists , or free variables.)

(1) All formulas in L^1 (A. Church, A. M.

Turning),

(2) $\exists x_1 \exists x_2 \dots \exists x_m \forall y_1 \forall y_2 \dots \forall y_n \mathfrak{A}$ (T. Skolem),

(3) $\exists x_1 \exists x_2 \exists x_3 \forall y_1 \forall y_2 \dots \forall y_n \exists z \mathfrak{A}$ (K. Gödel),

(4) $\exists x_1 \exists x_2 \forall y_1 \forall y_2 \dots \forall y_n \exists z \mathfrak{A}$ (L. Kalmar),

(5) $\exists x_1 \exists x_2 \forall y \exists z_1 \exists z_2 \dots \exists z_n \mathfrak{A}$ (J. Pepis),

(6) $\forall x \exists y \forall z \exists u_1 \exists u_2 \dots \exists u_n \mathfrak{A}$ (W. Ackermann),

(7) $\exists x_1 \exists x_2 \exists x_3 \forall y \mathfrak{A}$ or

$\exists x_1 \exists x_2 \forall y \exists z \mathfrak{A}$ (J. Surányi),

(8) $\exists x \forall y_1 \forall y_2 \exists z_1 \exists z_2 \mathfrak{A}$ or

$\forall x \exists y \forall z \exists u_1 \exists u_2 \mathfrak{A}$ (Surányi).

The decision problem has been solved affirmatively for the sets of formulas of the following forms, where it is assumed again that no function symbols appear and that \mathfrak{A} is as above.

(1) All formulas involving variables only on predicates with one argument. (L. Löwenheim, Skolem, H. Behmann),

(2) $\forall x_1 \forall x_2 \dots \forall x_m \mathfrak{A}$ (P. Bernays, M. Schönfinkel, Ackermann),

(3) $\forall x_1 \forall x_2 \dots \forall x_m \exists y_1 \exists y_2 \dots \exists y_n \mathfrak{A}$ (Bernays, Schönfinkel, Ackermann),

(4) $\forall x_1 \forall x_2 \dots \forall x_m \exists y_1 \exists y_2 \forall z_1 \forall z_2 \dots \forall z_n \mathfrak{A}$ (Gödel, Kalmar, K. Schütte).

(II) Results concerning formal systems on L^1 . Throughout the rest of this article we assume that no \dagger function variables appear. Predicate constants, function constants, and object constants may appear. By the decision problem for a formal system \mathfrak{Q} we mean the decision problem for all \dagger closed formulas in \mathfrak{Q} . Most of the results obtained so far concerning the decision problem for formal systems have been negative. Such results include those for formal systems formalizing natural number theory, the theory of rational integers, the elementary theory of \dagger groups, \dagger rings, \dagger fields, \dagger lattices, and the like, and axiomatic set theory (A. Tarski et al.).

The word problem for groups was solved negatively by P. S. Novikov (\rightarrow 161 Free Groups B). In connection with this decision problem, there are some investigations

by W. W. Boone, G. Higman, and others [14–17].

The decision problem for a formal system formalizing the elementary theory of \dagger Abelian groups has been solved affirmatively (W. Szmielew). Little is known about the decision problem concerning partial systems of formulas of given formal systems except the following: (1) the decision problem for the set of formulas of the form $\forall x_1 \forall x_2 \dots \forall x_m \mathcal{A}$ in a formal system (\rightarrow 161 Free Groups); (2) the Hilbert-type problem, which is the decision problem for the set of formulas of the form $\exists x_1 \exists x_2 \dots \exists x_m (t=s)$ in a formal system. In particular, the Hilbert-type problem in a formal system formalizing natural number theory is called Hilbert's tenth problem (\rightarrow 196 Hilbert). The latter is the problem of finding an algorithm for deciding whether a \dagger Diophantine equation has an integral solution.

This decision problem was studied by M. Davis, H. Putnam, J. Robinson, and others, and finally Yu. V. Matiyasevich solved it negatively by showing that every recursively enumerable relation is Diophantine [9]. (A relation $R(m_1, \dots, m_j)$ is called **Diophantine** if there is a polynomial $P(x_1, \dots, x_j, y_1, \dots, y_k)$ with integer coefficients such that $R(m_1, \dots, m_j)$ holds if and only if $P(m_1, \dots, m_j, y_1, \dots, y_k)=0$ has a solution for y_1, \dots, y_k in natural numbers.) In addition, some investigations have been made about the \dagger second-order predicate calculus L^2 , \dagger intuitionistic logic, etc. [1, 2].

References

- [1] W. Ackermann, Solvable cases of the decision problem, North-Holland, 1954.
- [2] S. C. Kleene, Introduction to metamathematics, Van Nostrand, 1952.
- [3] A. Tarski, Undecidable theories, North-Holland, 1953.
- [4] A. Church, An unsolvable problem of elementary number theory, Amer. J. Math., 58 (1936), 345–363.
- [5] S. C. Kleene and E. L. Post, The upper semi-lattice of degrees of recursive unsolvability, Ann. Math., (2) 59 (1954), 379–407.
- [6] G. E. Sacks, Degrees of unsolvability, Ann. Math. Studies, Princeton Univ. Press, 1963.
- [7] M. Davis, H. Putnam, and J. Robinson, The decision problem for exponential diophantine equations, Ann. Math., (2) 74 (1961), 425–436.
- [8] M. Davis and H. Putnam, Reductions of Hilbert's tenth problem, J. Symbolic Logic, 23 (1958), 183–187.
- [9] Yu. V. Matiyasevich, Diophantine enumerable sets (in Russian), Dokl. Akad. Nauk SSSR, 191 (1970), 279–282.

[10] H. Hermes, Enumerability, decidability, computability, Springer, 1965.

[11] H. Rogers, Theory of recursive functions and effective computability, McGraw-Hill, 1967.

[12] J. R. Shoenfield, Mathematical logic, Addison-Wesley, 1967.

[13] P. S. Novikov, On the algorithmic unsolvability of the word problem in groups (in Russian), Trudy Mat. Inst. Steklov., 44 (1955).

[14] G. Higman, Subgroups of finitely presented groups, Proc. Roy. Soc. London, (A) 262 (1961), 455–475.

[15] W. W. Boone, The word problem, Ann. Math., 77 (1963), 16–32.

[16] C. F. Miller, On group-theoretic decision problems and their classification, Ann. Math. Studies, Princeton Univ. Press, 1971.

[17] W. W. Boone, F. B. Cannonito, and R. C. Lyndon (eds.), Word problems, North-Holland, 1973.

98 (XXI.18) Dedekind, Julius Wilhelm Richard

Julius Wilhelm Richard Dedekind (October 6, 1831–February 12, 1916) was born in the city of Braunschweig in central Germany and studied at the University of Göttingen under C. F. \dagger Gauss, who was then in his later years. He received his doctorate at Göttingen with a thesis on the \dagger Euler integral. He was professor of mathematics from 1858 to 1862 at Zürich and from 1863 to 1894 at the Technische Hochschule in Braunschweig. During his early twenties, he wrote papers concerning analysis and the theory of probability, but in 1857 he began publishing papers on the theory of numbers. He edited \dagger Dirichlet's lectures on number theory (*Vorlesungen über Zahlen-theorie*, first edition 1863, fourth edition 1899) and concentrated on research in arithmetic and algebra. The theory of \dagger ideals, which he founded, was originally set out in a supplement (1863) to Dirichlet's *Vorlesungen*.

Dedekind treated subjects ranging from the axiomatic foundations of the theory of ideals to \dagger lattices and \dagger groups as algebraic systems. He was a pioneer of the abstract algebra of the 20th century. Among his notable achievements are the \dagger Dedekind zeta functions of \dagger algebraic number fields, \dagger Dedekind cuts in the theory of real numbers, the algebraic theory of \dagger algebraic functions (of which he was a coauthor with H. Weber), and the theory of natural numbers. He was one of the first to support

†Cantor's set theory. His theory of natural numbers was founded on the concept of sets and included the idea of †recursive functions.

References

- [1] R. Dedekind, *Gesammelte mathematische Werke I–III*, Vieweg, 1930–1932 (Chelsea, 1969).
 [2] M. Cantor and R. Dedekind, *Briefwechsel, Cantor–Dedekind*, *Actualités Sci. Ind.*, Hermann, 1937.

99 (IX.4) Degree of Mapping

A. Degree of Mapping

Let M^n and N^n be n -dimensional †closed †oriented † C^r manifolds (or †combinatorial manifolds). For example, $M^n = N^n = S^n$ (the † n -sphere). Their n th †homology groups with integral coefficients $H_n(M^n; \mathbf{Z})$ and $H_n(N^n; \mathbf{Z})$ are infinite cyclic groups generated by the fundamental homology classes $[M^n]$ and $[N^n]$, respectively (\rightarrow 201 Homology Theory). A continuous mapping $f: M^n \rightarrow N^n$ induces a homomorphism $f_*: H_n(M^n; \mathbf{Z}) \rightarrow H_n(N^n; \mathbf{Z})$, and there exists an integer d_f such that $f_*([M^n]) = d_f[N^n]$. This integer d_f is called the **degree of mapping** (or the **mapping degree**) of f . When $M^n = N^n$, d_f does not depend on the orientation of M^n .

If a continuous mapping $g: M^n \rightarrow N^n$ is †homotopic to f ($f \simeq g$), then we have $d_f = d_g$. If f is homotopic to a †constant mapping ($f \simeq 0$), then $d_f = 0$, while if f is a homeomorphism, then $d_f = \pm 1$. When $M^n = N^n$, a homeomorphism $f: M^n \rightarrow M^n$ is called an **orientation-preserving mapping** if $d_f = 1$ and an **orientation-reversing mapping** if $d_f = -1$.

Suppose that M^n and N^n are closed oriented n -dimensional combinatorial manifolds and that $f: M^n \rightarrow N^n$ is a †simplicial mapping. Let $\sum_i \sigma_i^n$ and $\sum_j s_j^n$ (σ_i^n, s_j^n are n -simplexes of M^n, N^n , respectively) represent $[M^n], [N^n]$, and let p_j (resp. q_j) be the number of n -simplexes σ_i^n such that $f(\sigma_i^n)$ is equal to s_j^n (resp. $-s_j^n$). Then $p_j - q_j$ is independent of the index j and equal to d_f .

Suppose that $f, g: S^n \rightarrow S^n$ are continuous mappings ($n \geq 1$). Then $f \simeq g$ if and only if $d_f = d_g$ (**Brouwer mapping theorem**). This implies that $\pi_n(S^n) \cong \mathbf{Z}$ (\rightarrow 202 Homotopy Theory).

B. Local Degree of Mapping

Suppose that M^n and N^n are n -dimensional oriented C^r (or combinatorial) manifolds and $f: M^n \rightarrow N^n$ is a continuous mapping. Suppose further that a point p of M^n has a neighborhood U such that $f(p) \neq f(q)$ for any point q contained in $U - \{p\}$. Then f induces a homomorphism $f_*: H_n(U, U - \{p\}) \rightarrow H_n(N, N - \{f(p)\})$ of n -dimensional †local homology groups with integral coefficients that are both isomorphic to \mathbf{Z} . If u and v are generators of the groups $H_n(U, U - \{p\})$ and $H_n(N, N - \{f(p)\})$ corresponding to orientations, respectively, then there exists an integer k such that $f_*(u) = kv$. We call this integer k the **local degree of mapping** f at p . If M^n and N^n are closed oriented C^r manifolds ($r \geq 1$) and $f: M^n \rightarrow N^n$ is a C^r mapping, then there exists a point r of N^n such that the set $f^{-1}(r)$ is a discrete subset $\{p_1, \dots, p_i\}$ of M^n , and each p_i has a neighborhood U_i satisfying the foregoing condition (†Sard's theorem). If k_i is the local degree of f at p_i , then $d_f = \sum k_i$.

C. Linking Numbers

Given two mutually disjoint smooth closed curves C_1 and C_2 in Euclidean 3-space, a quantity $\text{Lk}(C_1, C_2)$ indicating how closely they are interlinked with each other was given by Gauss as follows: Let C_i be expressed by the parameters $x_i = x_i(t_i)$ ($i = 1, 2$), where $x_i(t_i)$ are †continuously differentiable. Then the quantity

$$\text{Lk}(C_1, C_2) = -\frac{1}{4\pi} \int_{C_1} \int_{C_2} \frac{1}{|x_2 - x_1|^3} \times \det \begin{pmatrix} x_2 - x_1, \frac{dx_1}{dt_1}, \frac{dx_2}{dt_2} \end{pmatrix} dt_2 dt_1$$

is an integer called the **linking number** of C_1 and C_2 .

More generally, let M^n be an n -dimensional oriented †combinatorial manifold (or C^r manifold ($r \geq 1$)) and K and K^* its †cellular decompositions such that K^* is dual to K . Let z_1^r and z_2^s ($r + s = n - 1$) be †boundaries belonging to the complex K and K^* . Suppose that C^{r+1} is any †chain of K whose boundary is z_1^r . Then the †intersection number $[C^{r+1}] \cdot [z_2^s]$ does not depend on the choice of such a chain C^{r+1} . We set $\text{Lk}(z_1^r, z_2^s) = [C^{r+1}] \cdot [z_2^s]$ and call it the **linking number** of z_1^r and z_2^s . The **linking number** $\text{Lk}(\tilde{z}_1^r, \tilde{z}_2^s)$ of †singular boundaries $\tilde{z}_1^r, \tilde{z}_2^s$ ($r + s = n - 1$) of M^n is similarly defined by considering the approximations z_1^r, z_2^s of $\tilde{z}_1^r, \tilde{z}_2^s$ belonging to a suitable cellular decomposition K and its dual K^* . The number $\text{Lk}(\tilde{z}_1^r, \tilde{z}_2^s)$ is bilinear with respect to $\tilde{z}_1^r, \tilde{z}_2^s$, and we have $\text{Lk}(\tilde{z}_1^r, \tilde{z}_2^s) = (-1)^{r+s+1} \text{Lk}(\tilde{z}_2^s, \tilde{z}_1^r)$. In the example in 3-dimen-

sional Euclidean space \mathbf{R}^3 shown in the left half of Fig. 1, we have $\text{Lk}(\tilde{z}_1^1, \tilde{z}_2^2) = 1$, while $\text{Lk}(\tilde{z}_1^1, \tilde{z}_2^2) = 2$ for the example shown in the right half of the same figure. In particular, if \tilde{z}_1^1 is homologous to 0 in $M^n - |\tilde{z}_2^2|$, then we have $\text{Lk}(\tilde{z}_1^1, \tilde{z}_2^2) = 0$ (Fig. 2). Generally, if \tilde{z}_1^r and \tilde{z}_1^r are homologous in $M^n - |\tilde{z}_2^2|$, then $\text{Lk}(\tilde{z}_1^r, \tilde{z}_2^2) = \text{Lk}(\tilde{z}_1^r, \tilde{z}_2^2)$.

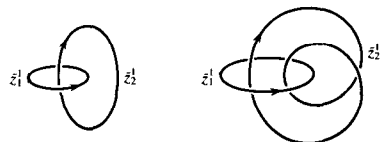


Fig. 1

D. Order of a Point with Respect to a Cycle

Let M^n be an n -dimensional oriented combinatorial manifold (or C^r manifold ($r \geq 1$)) with the n th Betti number $b_n = 0$, \tilde{z}^{n-1} an $(n-1)$ -dimensional singular boundary of M^n , and o a point of M^n that is not contained in $|\tilde{z}^{n-1}|$. We set $\text{ord}(\tilde{z}^{n-1}, o) = \text{Lk}(\tilde{z}^{n-1}, o)$ and call it the **order of the point** o with respect to \tilde{z}^{n-1} . For example, when $M^n = \mathbf{R}^2$ and $\tilde{z}^1 = \{f(t) | 0 \leq t \leq 1, f(0) = f(1)\}$, where f is a continuous function, the order $\text{ord}(\tilde{z}^1, o)$ is equal to the **rotation number** around o of a moving vector $\overrightarrow{of}(t)$ as t varies from 0 to 1. This $\text{ord}(\tilde{z}^1, o)$ stays invariant as the point o moves in a * connected component of the complement $\mathbf{R}^2 - |\tilde{z}^1|$ (Fig. 3). On the other hand, if $\tilde{z}_i^1 = \{f_i(t) | 0 \leq t \leq 1, f_i(0) = f_i(1)\}$ ($i = 0, 1$) are closed curves in \mathbf{R}^2 and the distance $\rho(f_0(t), f_1(t))$ is smaller than $\rho(f_0(t), o)$ for all t in the interval $[0, 1]$, then we have $\text{ord}(\tilde{z}_0^1, o) = \text{ord}(\tilde{z}_1^1, o)$ (**Rouché's theorem**).

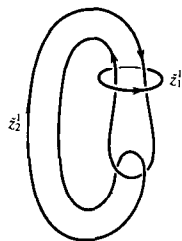


Fig. 2

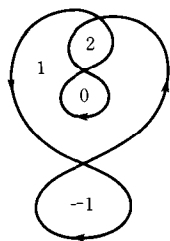


Fig. 3

References

- [1] P. S. Aleksandrov and H. Hopf, *Topologie I*, Springer, 1935 (Chelsea, 1965).
- [2] H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Teubner, 1934 (Chelsea, 1968); English translation, *A textbook of topology*, Academic Press, 1980.
- [3] J. W. Milnor, *Topology from the differentiable viewpoint*, Univ. Press of Virginia, 1965.

100 (X.17) Denjoy Integrals

A. History

For a real-valued function $f(x)$ of a real variable to be * Lebesgue integrable, it is necessary and sufficient that there exist an * absolutely continuous function $F(x)$ such that $F'(x) = f(x)$ at * almost all points x (\rightarrow 221 Integration Theory D). In general, the derivative of a function is not necessarily Lebesgue integrable. A function $f(x)$ is Lebesgue integrable if and only if $|f(x)|$ is integrable. Hence a function which has an improper Riemann integral is not necessarily Lebesgue integrable (\rightarrow 221 Integration Theory A). For this reason, it is desirable to extend the concept of Lebesgue integrals. In 1912, A. Denjoy constructively defined a new concept of integrals (Denjoy integral in the restricted sense; \rightarrow Section D), which is an extension of both Lebesgue and Riemann integrals. Later, N. N. Luzin provided the descriptive theory of this integral. Independently, and nearly simultaneously, A. J. Khinchin and Denjoy defined a more general integration (Denjoy integral in the wide sense (1916); \rightarrow Section D).

In 1914, O. Perron, independently of Denjoy, defined a concept of integrals (Perron integrals) that is equivalent to that of Denjoy integrals in the restricted sense. To establish this concept, Perron considered the differential equation $y' = f(x)$ and utilized a method similar to the one used in the proof of the existence theorem for the solution of the differential equation $y' = f(x, y)$. However, the concept of Denjoy integrals is inadequate to treat unbounded functions. Thus to extend the concepts of Riemann and Lebesgue integrals, various ideas have been introduced; for example, Denjoy (1921), J. C. Burkill (1951), and R. D. James (1950) introduced new concepts as byproducts of investigations concerning the coefficients of trigonometric series [2, 3]. The A -integral concept devised by A. N. Kolmogorov was meant to deal with the problem of the conjugate function of Fourier series [4]. As a certain completion of the space of functionals of step functions, K. Kunugui defined the notion of E. R. integrals (1956), which coincides with that of A -integrals in a special case [5, 6].

What has been stated so far deals only with functions of a real variable. Concerning the extension of Denjoy integrals to the case of several variables, research has been done by M. Loomis, S. Kempisty, S. Nakanishi (née Enomoto), and others [7, 8].

B. Approximate Derivative

If we have $\lim_{h \rightarrow 0, k \rightarrow 0} m\{E \cap (\xi - h, \xi + k)\} / (h + k) = 1$ at a point ξ of a measurable subset E of the real line, where m is the 1-dimensional Lebesgue measure, the point ξ is called a **point of density** for E . Almost all the points of E are points of density for E (**Lebesgue's density theorem**). Let E be a measurable set having x_0 as a point of density, and let $F(x)$ be a measurable function on E . If there exists a number l such that for each $\varepsilon > 0$, x_0 is a point of density for the set $\{x | l - \varepsilon \leq (F(x) - F(x_0)) / (x - x_0) \leq l + \varepsilon, x \in E\}$, then l is called the **approximate derivative** of $F(x)$ at x_0 and is denoted by $ADF(x_0)$. If $ADF(x_0)$ exists, $F(x)$ is said to be **approximately derivable** at x_0 . If $F(x)$ is approximately derivable at each point of E , then $F(x)$ is said to be **approximately derivable in E** . If $F'(x)$ exists at a point x , then $ADF(x)$ exists at x , and we have $ADF(x) = F'(x)$. However, there exists a continuous function $F(x)$ that is approximately derivable at almost all points of an interval and yet not differentiable at any point of a set of positive measure.

C. Generalized Absolute Continuity

Let E be a set in \mathbf{R} , and let $F(x)$ be a real-valued function whose domain contains E . If for each $\varepsilon > 0$, there is a $\delta > 0$ such that for every sequence $\{[a_n, b_n]\}$ of nonoverlapping intervals whose endpoints belong to E the inequality $\sum(b_n - a_n) < \delta$ implies $\sum|F(b_n) - F(a_n)| < \varepsilon$, then the function $F(x)$ is said to be **absolutely continuous** on E . We denote by AC the set of all functions that are absolutely continuous on E . If $F(x)$ is continuous on E and E is the union of a countable sequence of sets E_n on each of which $F \in AC$, then $F(x)$ is called a **generalized absolutely continuous function**, and we write $F \in GAC$. If $F \in GAC$, $ADF(x)$ exists almost everywhere.

If, for each $\varepsilon > 0$, there is a $\delta > 0$ such that for every sequence $\{[a_n, b_n]\}$ of nonoverlapping intervals whose endpoints belong to E the inequality $\sum(b_n - a_n) < \varepsilon$ implies $\sum O\{F; [a_n, b_n]\} < \varepsilon$ ($O\{F; [a_n, b_n]\}$ denotes the oscillation of the function $F(x)$ in $[a_n, b_n]$, i.e., the difference between the least upper bound and the greatest lower bound of the values assumed by $F(x)$ on $[a_n, b_n]$), then $F(x)$ is said to be **absolutely continuous in the restricted sense** (or **absolutely continuous (*)**) on E ; and we write $F \in AC(*)$. Just as we defined the notions of generalized absolute continuity and absolute continuity, so we define the notions of **generalized absolute continuity in the restricted sense** and **generalized absolute continuity (*)**. Thus $F \in GAC(*)$ means that $F(x)$ is a generalized ab-

solute continuous (*) function on E . If $F \in GAC(*)$, then $F'(x)$ exists almost everywhere.

D. Definitions of Denjoy Integrals

Let $f(x)$ be a real-valued function defined on $I = [a, b]$. If for $f(x)$ there exists a function $F(x)$ that belongs to GAC on I and for which $ADF(x) = f(x)$ holds almost everywhere, then $f(x)$ is said to be **Denjoy integrable in the wide sense** (or **D-integrable**) on I . We call $F(b) - F(a)$ the **definite D-integral** of $f(x)$ over I , and denote the value by $(D) \int_a^b f(x) dx$. The function $F(x)$ is called an **indefinite D-integral** of $f(x)$ on I . Similarly, we obtain the definition of **Denjoy integral in the restricted sense** (or **D(*)-integral**) by replacing GAC by $GAC(*)$ and $ADF(x)$ by $F'(x)$ in the definition of the D-integral. If a continuous function $F(x)$ satisfies the equality $ADF(x) = f(x) \neq \pm \infty$ ($F'(x) = f(x) \neq \pm \infty$) for all except countably many points in I , then $F(x)$ is an **indefinite D-integral** (D(*)-integral) of $f(x)$. A Lebesgue-integrable function is D(*)-integrable, a D(*)-integrable function is D-integrable, and a D-integrable function that is almost everywhere nonnegative is Lebesgue integrable.

E. Constructive Definition of Integrals

Let S be a functional whose domain $\bigcup_I K(S; I)$ consists of the union of sets $K(S; I)$ of real-valued functions defined on closed intervals $I = [a, b]$. If f belongs to $K(S; I)$, we denote the value $S(f)$ by $S(f; I)$. Such a functional S is called an **integral operator** if the following three conditions are satisfied: (1) If $f \in K(S; I_0)$ and I is an arbitrary interval contained in I_0 , then the restriction f_I to I of f also belongs to $K(S; I)$. Also, $S(f; I)$ is a continuous additive function of the interval $I \subset I_0$. (2) Let $I_1 = [a, b]$, $I_2 = [b, c]$, and $I = [a, c]$ ($a < b < c$). If for a function f defined on I , $f_1 \in K(S; I_1)$ and $f_2 \in K(S; I_2)$, where $f_1 = f_{I_1}$ and $f_2 = f_{I_2}$, then $f \in K(S; I)$. (3) If f is identically 0 on I , then $f \in K(S; I)$ and $S(f; I) = 0$. For two integral operators S_1 and S_2 , we say that S_2 includes S_1 (or S_1 is weaker than S_2) if $K(S_1; I) \subset K(S_2; I)$ for every I and $S_1(f; I) = S_2(f; I)$ for every $f \in K(S_1; I)$. The D-integral (D(*)-integral) is the weakest integral operator containing the Lebesgue integral and satisfying the following two conditions, (C) and (H) (resp. H(*)): (C) **Cauchy's condition**. If, for every function f defined on I_0 , we have $f_I \in K(S; I)$ for any $I = [a + \delta, b - \varepsilon] \subset I_0 = [a, b]$, and also if the finite limit $\lim_{\delta \rightarrow 0, \varepsilon \rightarrow 0} S(f; I)$ exists, then $f \in K(S; I_0)$ and $S(f; I_0)$ coincides with the foregoing limit value. (H) **Harnack's condition**. Let E be a closed subset of I_0 , $\{I_k\}$ be a sequence of inter-

vals contiguous to the set consisting of the points of E and the endpoints of I_0 , and f be a function on I_0 satisfying the following three conditions: (i) $f_E \in K(S; I_0)$, where $f_E(x) = f(x)$ whenever $x \in E$ and $f_E = 0$ otherwise; (ii) $f_k = f_{I_k} \in K(S; I_k)$ for each k ; and (iii) $\sum_k |S(f_k; I_k)| < +\infty$ and $\lim_{k \rightarrow \infty} O(S; f_k; I_k) = 0$ when the sequence $\{I_k\}$ is infinite. Then it follows that $f \in K(S; I_0)$ and $S(f; I_0) = S(f_E; I_0) + \sum_k S(f_k; I_k)$. (Here $O(S; f_k; I_k)$ denotes the **variation** of $S(f_k)$ on I_k , that is, the least upper bound of the numbers $|S(f_j; J)|$, where J denotes any subinterval of I_k .) We obtain condition (H*) by replacing condition (iii) in (H) with a more restrictive condition: $\sum_k O(S; f_k; I_k) < +\infty$. The constructive definition of the Denjoy integral in the wide sense (the Denjoy integral in the restricted sense) is obtained by †transfinite induction starting with the Lebesgue integral and using two methods, (C) and (H) (resp. (H*)), of extensions.

F. Perron Integrals

Given a function $f(x)$ defined on an interval $[a, b]$, suppose that $F(x)$ is a function defined on the same interval such that (1) $F(x) \geq f(x)$; (2) $F(x) \neq -\infty$ (resp. (1') $\bar{F}(x) \leq f(x)$; (2') $\bar{F}(x) \neq +\infty$) at every point x , where $\underline{F}(x)$ (resp. $\bar{F}(x)$) denotes the †lower (upper) derivative of $F(x)$. In this case, $F(x)$ is called a **major (minor) function** of $f(x)$. If for any $\varepsilon > 0$ there is a major function $\psi(x)$ and a minor function $\varphi(x)$ of $f(x)$ such that $\psi(b) - \varphi(b) < \varepsilon$, then $f(x)$ is said to be **Perron integrable**. We denote by $(P) \int_a^b f(x) dx$ the value $\inf_{\psi} \{\psi(b) - \psi(a)\} = \sup_{\varphi} \{\varphi(b) - \varphi(a)\}$.

G. Properties of Integrals

If $\{f_n\}$ is a nondecreasing sequence of functions that are D-integrable on an interval $[a, b]$ and whose D-integrals over $[a, b]$ constitute a sequence bounded from above, then the function $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ is D-integrable on $[a, b]$, and we have

$$(D) \int_a^b f(x) dx = \lim_{n \rightarrow \infty} (D) \int_a^b f_n(x) dx.$$

If $F(x)$ is a function of †bounded variation and $g(x)$ is a D-integrable function on an interval $[a, b]$, then $F(x)g(x)$ is D-integrable on $[a, b]$; moreover, denoting by $G(x)$ the indefinite D-integral of $g(x)$, the following formula is valid:

$$(D) \int_a^b F(x)g(x) dx = G(b)F(b) - G(a)F(a) - \int_a^b G(x)dF(x),$$

where the last term is the †Stieltjes integral (**integration by parts**).

If $F(x)$ is a nondecreasing function and $g(x)$ is D-integrable on $[a, b]$, there is a point ξ in $[a, b]$ for which the following formula is valid:

$$(D) \int_a^b g(x)F(x) dx = F(a) \cdot (D) \int_a^{\xi} g(x) dx + F(b) \cdot (D) \int_{\xi}^b g(x) dx$$

(the **second mean value theorem**).

The foregoing theorems remain valid if D is replaced by D^* in the hypotheses and conclusions.

References

- [1] S. Saks, *Theory of the integral*, Warsaw, 1937.
- [2] R. D. James, *Integrals and summable trigonometric series*, Bull. Amer. Math. Soc., 61 (1955), 1–15.
- [3] R. Henstock, *Theory of integration*, Butterworth, 1963.
- [4] Yu. S. Ochan, *Generalized integrals* (in Russian), Mat. Sb., 28 (70), no. 2 (1951), 293–336.
- [5] K. Kunugui, *Sur une généralisation de l'intégrale*, Fundamental and Applied Aspects of Math., 1 (1959), 1–30.
- [6] H. Okano, *Sur une généralisation de l'intégral* (E. R.) et un théorème général de l'intégration par parties, J. Math. Soc. Japan, 14 (1962), 430–442.
- [7] S. Kempisty, *Sur les fonctions absolument continue d'intervalle*, Fund. Math., 27 (1936), 10–37.
- [8] S. Enomoto, *Sur une totalisation dans les espaces de plusieurs dimensions I, II*, Osaka Math. J., 7 (1955), 69–102, 157–187.

101 (XXI.19) Descartes, René

René Descartes, (March 31, 1596–February 11, 1650), philosopher, mathematician, and natural scientist, was born in the province of Touraine in France. He became dissatisfied with his studies of scholastic philosophy in the Jesuit Academy in La Flèche, and later, in 1619, while stationed in Ulm during a tour of duty in the army, he underwent a philosophical conversion. He had an idea of methodologically unifying the various fields of interest to him using mathematics as a model. He returned to Paris in 1621, but moved to Holland in 1628 to concentrate on his work. Swe-

den's Queen Christina invited him in 1649 to that country, where he died the next year, evidently from a combination of cold and overwork.

Descartes, often considered the founder of modern philosophy, discarded early the traditional theological view of the world and stated that all knowledge should be recognized as logical only after it has been submitted to rational criticism. This ushered in the modern view of the world based on mathematics and physics. In 1637, he published *Géométrie* as an appendix to his *Discours de la méthode*, which also contained his works on optics and meteorology. In it, he promoted F. Viète's symbolic algebra, which he applied to geometric problems. His idea that algebra could be used as a general method for geometry established him as the founder of analytic geometry.

References

- [1] R. Descartes, Oeuvres I–XII, C. Adam, P. Tannery, and Léopold Cerf (eds.), 1897–1910.
[2] H. Lefèvre, Descartes, Paris, 1947.

102 (XVIII.12) Design of Experiments

A. General Remarks

The **design of experiments** is a part of the **statistical planning** required to collect the data appropriate to the purpose of statistical inference (→ 401 Statistical Inference) in various fields of scientific research and application. The main purposes of the design of experiments are (1) to analyze a given statistical linear model (→ 403 Statistical Models) and (2) to devise a “good” statistical linear model. Sometimes this term also refers to a statistical method including the analysis of variances. Thus the purpose of designing an experiment is to provide the most efficient and economical methods of reaching valid and relevant conclusions from that experiment (→ 403 Statistical Models).

R. A. Fisher, whose contributions to statistical theory were remarkable and far-ranging, propounded three required principles to control the experimental field in order to guarantee the validity of statistical methods and to increase the sensitivity of experiments: (i) **replication**, or the repetition of the set of all the phenomena to be compared in the experiment, for the evaluation of experimental error variance; (ii) **randomization**, or the procedure

allotting various experimental treatments at random, in order to change systematic errors into random errors; and (iii) **local control**, or the procedure which makes the variation within each experimental block as small as possible, in order to minimize or remove systematic errors. These are called **Fisher's three principles**. A design satisfying principles (i) and (ii) is called a **completely randomized design**; it enables us to attach a probability statement to estimated treatment differences by obtaining valid estimate of experimental error variance.

Let an n -dimensional random variable $\mathbf{Y} = (Y_1, \dots, Y_n)'$ be represented by a linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\xi} + \mathbf{W}, \quad (1)$$

where \mathbf{X} is a given $n \times s$ real matrix, $\boldsymbol{\xi} = (\xi_1, \dots, \xi_s)'$ is an s -vector, and $\mathbf{W} = (W_1, \dots, W_n)'$ is a random vector with the expectation $E(\mathbf{W}) = \mathbf{0}$. Then \mathbf{Y} is called the **observation vector**, \mathbf{W} the **error vector**, $\boldsymbol{\xi}$ the **effect vector** of \mathbf{Y} , and \mathbf{X} the **design matrix**.

According to the properties of the effect vector $\boldsymbol{\xi}$, the linear model (1) is separated into three classes: (i) The class of **fixed-effects models** for which $\boldsymbol{\xi}$ is a fixed unknown parameter. In this case, the component ξ_i of $\boldsymbol{\xi}$ is called a **fixed effect**, and a linear function $\pi = \mathbf{F}'\boldsymbol{\xi}$ of $\boldsymbol{\xi}$, with a given coefficient vector \mathbf{F} is called a **linear parameter** or **parametric function**. (ii) The class of **random-effects models** for which the components ξ_i of $\boldsymbol{\xi}$ are random variables. In this case, each component ξ_i is called a **random effect**, and $\boldsymbol{\xi}$ is denoted by $\boldsymbol{\Xi}$. (iii) The class of **mixed models** for which there are both fixed effects ξ_i and random effects ξ_j in $\boldsymbol{\xi}$. In this case, the model (1) becomes

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\xi}^1 + \mathbf{X}_2\boldsymbol{\Xi}^2 + \mathbf{W}, \quad (2)$$

where $\boldsymbol{\xi}^1 = (\xi_1, \dots, \xi_r)'$ is a fixed-effect vector and $\boldsymbol{\Xi}^2 = (\xi_{r+1}, \dots, \xi_s)'$ is a random-effect vector. The conditions frequently assumed for the distribution law of \mathbf{Y} are: (a) The errors W_i ($i = 1, \dots, n$) are uncorrelated and $E(W_i) = \dots = E(W_n) = 0$. (b) The errors W_i ($i = 1, \dots, n$) have a common unknown variance σ^2 . (c) The errors W_i ($i = 1, \dots, n$) have the normal distribution. (d) The random effects ξ_j are uncorrelated and independent of the error vector \mathbf{W} . (e) The random effects ξ_j have a common unknown variance σ_j^2 . (f) The random effects ξ_j have the normal distribution.

Let $L(\mathbf{X})$ be a linear subspace of \mathbf{R}^n spanned by the column vectors of \mathbf{X} . The linear model

$$\mathbf{Y} = \tilde{\mathbf{X}}\boldsymbol{\xi} + \mathbf{W} \quad (3)$$

is called a hypothesis on the linear model (1) if $L(\tilde{\mathbf{X}}) \subset L(\mathbf{X})$.

The main issues of the theory of design of experiments are concerned with (I) statis-

tical inferences, such as estimation or testing hypotheses, under models (1), (2), (3), (i), (ii), (iii); (II) determination of the design matrix X satisfying certain optimal conditions; (III) construction of a theoretical foundation that can explain the validity of the statistical treatment of the observed data by means of the above models.

B. Block Design

The design of experiments is described here in terms of the so-called **block design**. There are n experimental units $\alpha = 1, \dots, n$ called **plots**, and an observation Y_α is assigned to each plot α . A **block** is constructed with several plots under Fisher's principle (iii), and the number of plots in a block is called the **block size**, the j th one being denoted by k_j , $j = 1, \dots, b$, with $\sum_j k_j = n$. One of v operations, called **treatments or varieties**, is applied to each plot. It is assumed that the observation Y_α at the plot α in the j th block under the i th treatment has the structure

$$Y_\alpha = \xi_i + \eta_j + W_\alpha.$$

The ξ_i , $i = 1, \dots, v$, are called **treatment effects**, and the η_j , $j = 1, \dots, b$, **block effects**. It is also assumed that $\sum_i \xi_i = 0$. In this case, Y is represented in matrix notation as

$$Y = \Phi\xi + \Psi\eta + W, \quad (4)$$

where $\Phi = (\varphi_{\alpha i})$, $\alpha = 1, \dots, n$, $i = 1, \dots, v$, with

$$\varphi_{\alpha i} = \begin{cases} 1 & \text{when the } i\text{th treatment is applied} \\ & \text{to the plot } \alpha, \\ 0 & \text{otherwise,} \end{cases}$$

and $\Psi = (\psi_{\alpha j})$, $\alpha = 1, \dots, n$, $j = 1, \dots, b$, with

$$\psi_{\alpha j} = \begin{cases} 1 & \text{when the plot } \alpha \text{ belongs to} \\ & \text{the } j\text{th block,} \\ 0 & \text{otherwise.} \end{cases}$$

Here it is assumed that $\sum_i \varphi_{\alpha i} = 1$, $\sum_\alpha \varphi_{\alpha i} = r_i \geq 1$, $\sum_i r_i = n$, $\sum_j \psi_{\alpha j} = 1$, and $\sum_\alpha \psi_{\alpha j} = k_j \geq 1$. We call r_i the **number of replications** of the i th treatment. We set $N = (n_{ij}) = \Phi'\Psi$. Then n_{ij} is the number of observations in the j th block to which the i th treatment is applied. The matrix N is called the **incidence matrix** of the block design.

In any experiment, each plot has its own effect. The blocks are constructed so that this plot effect in each block becomes as homogeneous as possible, although it is impossible to eliminate the effect completely. For this purpose, randomization is adopted. Suppose that we are given k plots in a block and v treatments ($k \leq v$). Then randomization is

utilized to select a treatment out of v treatments to be allocated to each plot so that the selection is "at random." Then the plot effects are random, and the error term in model (4) can be considered to be the sum of a plot effect and an original error.

A block design satisfying Fisher's three principles is called a **randomized block design**. Blocks that can accommodate all the treatments to be studied are called **complete blocks**. Those that do not contain all the treatments are called **incomplete blocks**. Blocking can be considered to be an extension of pairing. In the terminology of block design, we can say that many experiments are block experiments, or even (tautologically, if we allow those with only one block) that all experiments are block experiments.

In a block design $N = (n_{ij})$, treatments i_0 and i_l are said to be **connected** if there exists a chain

$$i_0 j_1 i_1 j_2 \dots i_{l-1} j_l i_l$$

of integers such that $1 \leq i_p \leq v$ ($p = 0, 1, 2, \dots, l$), $1 \leq j_q \leq b$ ($q = 1, 2, \dots, l$), and $n_{i_0 j_1} > 0$, $n_{i_1 j_2} > 0, \dots, n_{i_{l-1} j_l} > 0$, $n_{i_l j_l} > 0$. If all pairs of treatments are mutually connected, then the design is said to be **connected**. In this case, the rank of the matrix C defined in Section C is $v - 1$. If the design is disconnected, then the incidence matrix N can be partitioned into two or more connected portions, e.g.,

$$N = \begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix}.$$

Thus without loss of generality we can restrict ourselves to the connected case.

C. Estimation under the Fixed-Effects Model

Conditions (a) and (b) of Section A are assumed here. The normal equation which gives the least square estimates $\hat{\xi}$ and $\hat{\eta}$ of ξ and η , respectively, is

$$\begin{pmatrix} \Phi' \\ \Psi' \end{pmatrix} (\Phi, \Psi) \begin{pmatrix} \hat{\xi} \\ \hat{\eta} \end{pmatrix} = \begin{pmatrix} \Phi' \\ \Psi' \end{pmatrix} Y. \quad (5)$$

Set $\Phi'\Phi = \text{diag}(r_1, \dots, r_v) = D_r$, $\Psi'\Psi = \text{diag}(k_1, \dots, k_b) = D_k$, $C = D_r - N D_k^{-1} N'$, $Q = (\Phi' - N D_k^{-1} \Psi') Y$, where $\text{diag}(\dots)$ means a diagonal matrix with the diagonal elements \dots . Then (5) reduces to

$$C\hat{\xi} = Q, \quad \hat{\eta} = D_k^{-1} (\Psi' Y - N' \hat{\xi}). \quad (6)$$

Let L be an orthogonal matrix that transforms the matrix C to a diagonal form; that is, $L'CL = \text{diag}(\rho_1, \dots, \rho_{v-1}, 0) = \Lambda$, $\rho_i > 0$ for all i . Set $\Lambda^* = \text{diag}(\rho_1^{-1}, \dots, \rho_{v-1}^{-1}, 0)$, $C^* = L\Lambda^*L'$. Then $\hat{\xi}$

$= C^*Q$ is a particular solution of (6) and $\sum_i \hat{\xi}_i = 0$. A parametric function $\pi = F'\xi$ with coefficient vector $F = (F_1, F_2, \dots, F_v)'$ is called a **treatment contrast** if the sum $\sum_i F_i$ of coefficients vanishes. A treatment contrast $\pi = F'\xi$ is called a **normalized contrast** if $F'F = 1$. It is called the **elementary contrast** if F has only two nonzero elements 1 and -1 . Elementary contrasts of treatment effects show the comparison of treatments involved in them. When a design is connected, any contrast π is estimable, and the best linear unbiased estimate of π is $\hat{\pi} = F'\hat{\xi}$. Furthermore, if f_i is the eigenvector of the matrix C with unit length corresponding to an eigenvalue ρ_i and $F = \sum_i a_i f_i$, then the variance of the estimate $\hat{\pi}$ is given by $\sigma^2 \sum_i a_i^2 / \rho_i$. The following properties are equivalent: (i) A design is connected. (ii) Any treatment contrast is estimable. (iii) The rank of the matrix C is $v-1$. (iv) The minimum eigenvalue 0 of $D_r^{-1/2} C D_r^{-1/2}$ is simple and other eigenvalues θ_1 satisfy $0 < \theta_1 \leq 1$. (v) The maximum eigenvalue 1 of $D_r^{-1/2} N D_k^{-1} N' D_r^{-1/2}$ is simple and other eigenvalues θ_2 satisfy $0 \leq \theta_2 < 1$. (vi) There exists a positive integer p such that each element of $(D_r^{-1/2} N D_k^{-1} N' D_r^{-1/2})^p$ is positive. It holds that $\theta_2 = 1 - \theta_1$.

D. Test of a Hypothesis $H: \xi_1 = \dots = \xi_v$ in the Fixed-Effects Model

Conditions (a), (b), and (c) of Section A are assumed here. The hypothesis $H: \xi_1 = \dots = \xi_v$ is represented by

$$Y = \Psi\eta + W, \quad (7)$$

Consider a direct sum decomposition $R^n = L(\Gamma) + L_F^\perp(\Psi) + L_\Psi^\perp(\Phi, \Psi) + L_{\Phi, \Psi}^\perp$ of R^n , where $L_B^\perp(A)$ and L_B^\perp stand for the orthocomplements of $L(B)$ with respect to $L(A)$ and R^n , respectively, and $\Gamma = (1, 1, \dots, 1)' \in R^n$. The projection operator matrices for the decomposed subspaces $L(\Gamma)$, $L_F^\perp(\Psi)$, $L_\Psi^\perp(\Phi, \Psi)$, and $L_{\Phi, \Psi}^\perp$ are denoted by P_1 , P_2 , P_3 , and P_4 , respectively. Then we have

$$\begin{aligned} P_1 &= n^{-1} E_{nn}, & P_2 &= \Psi D_k^{-1} \Psi' - n^{-1} E_{nn}, \\ P_3 &= (I_n - \Psi D_k^{-1} \Psi') \Phi C^* \Phi' (I_n - \Psi D_k^{-1} \Psi'), \\ P_4 &= I_n - P_1 - P_2 - P_3, \end{aligned}$$

where E_{ab} is an $a \times b$ matrix whose entries are all unity and I_n is the $n \times n$ identity matrix.

The analysis of variance for the hypothesis (7) in model (4) is given by

$$Y'Y = Y'P_1 Y + Y'P_2 Y + Y'P_3 Y + Y'P_4 Y.$$

This is called an **intra-block analysis**. A usual test for the hypothesis H is given by a critical

region with

$$F = \frac{n-v-b+1}{v-1} \frac{Y'P_3 Y}{Y'P_4 Y} > \text{constant}$$

(\rightarrow 403 Statistical Models).

E. Optimal Block Design

A block design is said to be **optimal** when it minimizes the variance of the estimate $\hat{\pi}$ of a normalized contrast π . Suppose that the number v of treatments, the number b of blocks, and each block size k_j , $j = 1, \dots, b$, are given. Under each of the following criteria the corresponding block design is optimal in the sense indicated: For positive eigenvalues, $\rho_1, \rho_2, \dots, \rho_{v-1}$, of the matrix C in Section C, (I) $\prod_{i=1}^{v-1} \rho_i$ is maximal (**D-optimality**); (II) $\min \rho_i$ is maximal (**E-optimality**); (III) $\sum_{i=1}^{v-1} \rho_i^{-1}$ is minimal (or the average variance of the estimates of all normalized contrasts of the parameters ξ_i is minimal) (**A-optimality**).

If $\rho_1 = \dots = \rho_{v-1} = (n-b)/(v-1) (= \rho, \text{ say})$ and n_{ij} is either 1 or 0, then the design is optimal for each of the optimality criteria (I), (II), and (III). In this case, we have

$$C = \rho(I_v - v^{-1} E_{vv}).$$

Such a design is called a **variance-balanced block design** in which every normalized contrast is estimable with the same variance. If $D_r^{-1} N D_k^{-1} N' - n^{-1} E_{v1} r' = \mu(I_v - n^{-1} E_{v1} r')$ (or $C = (1 - \mu)(D_r - n^{-1} r r')$), every normalized contrast is estimable with the same efficiency $1 - \mu$, where $r = D_r E_{v1}$ and $0 < \mu < 1$. Such a design is called an **efficiency-balanced block design**. For a block design, any two of the following properties imply the third: (i) The design is variance-balanced; (ii) the design is efficiency-balanced; (iii) the design is equireplicated.

When all block sizes k_j equal some number k independent of j , all numbers r_i of replications equal some number r independent of i , and $\lambda_{ii'} = \sum_j n_{ij} n_{i'j}$ (number of times that both treatments i and i' are applied to the same block) equals some number λ independent of i and i' , then the design is combinatorially balanced. The design is usually called the **balanced incomplete block design (BIBD)** if these three conditions are fulfilled and $k < v$. A BIBD is both variance-balanced and efficiency-balanced. If all the treatments are replicated the same number of times and the blocks are of the same size in a block design, then the only variance-balanced design is a BIBD, provided such a design exists. A BIBD is often denoted by $\text{BIBD}(v, b, r, k, \lambda)$ and we have the relations $vr = bk$, $\lambda(v-1) = r(k-1)$ and $v \leq b$, among the parameters; the last

relation is called **Fisher's inequality**. The design is said to be **symmetric** when $v=b$. Furthermore, if $v=b$ is even, then $r-\lambda$ must be a perfect square. If $v=b$ is odd, then the equation $x^2=(r-\lambda)y^2+(-1)^{(v-1)/2}\lambda z^2$ must have a solution in integers x, y, z not all zero. This is called the **Bruck-Ryser-Chowla theorem**. Necessary conditions for the existence of a BIBD have been obtained. One of these conditions is stated in terms of the Hasse-Minkowski p -invariant $C_p(A)=(-1, -1)_p \prod_{i=1}^n (D_i, -D_{i-1})_p$, where $(m, m')_p$ is the \dagger Hilbert norm-residue symbol, $D_0=1$, and D_i is the principal minor of the $n \times n$ matrix A . Another is described in terms of the embedding of a quasiresidual design into the corresponding symmetric BIBD. No effective necessary and sufficient condition for the existence of a BIBD has been obtained. In general, it is conjectured that for a positive integer k , with finitely many exceptions, BIBD(v, b, r, k, λ)s exist for all pairs v, λ of positive integers satisfying $v > k$, $\lambda(v-1) \equiv 0 \pmod{k-1}$ and $\lambda v(v-1) \equiv 0 \pmod{k(k-1)}$.

The known methods of constructing block designs are of two main types, direct and recursive. Recursive methods are a way of constructing designs from smaller ones. Direct methods yield easier constructions, but are applicable only for special values of the parameters. A direct method of constructing a BIBD designates the treatments and blocks, respectively, as the points and subspaces of the \dagger projective space and the \dagger affine space over a \dagger finite field. To explain another method of constructing a BIBD, we let G be an \dagger additive group of order n and $x_1^{(i)}, \dots, x_m^{(i)}$ be m treatments corresponding to each element $x^{(i)}$ of the group ($i=1, 2, \dots, n$). The treatment $x_\alpha^{(i)}$ is said to belong to the α th class ($\alpha=1, 2, \dots, m$), and a pair $(x_\alpha^{(i)}, x_\beta^{(j)})$ of treatments is called a **difference** of type $(\alpha, \beta, x^{(p)})$ if $x^{(i)} - x^{(j)} = x^{(p)} (\neq 0)$. We can form t blocks of size k

$$B_1 = \{x_{\alpha_1}^{(i_1)}, \dots, x_{\alpha_k}^{(i_k)}\}, \quad \dots, \\ B_t = \{x_{\beta_1}^{(j_1)}, \dots, x_{\beta_k}^{(j_k)}\},$$

such that each block B_i contains exactly r treatments belonging to the α th class ($\alpha=1, \dots, m$) and among all pairs of treatments in the same block there are λ differences of each type $(\alpha, \beta, x^{(p)})$. Such a set of t blocks is called a **difference set**. The t blocks in a difference set are called **initial blocks**. Given such a difference set, we can obtain nt blocks by joining elements of G to the elements of each B_s ($s=1, 2, \dots, t$). These nt blocks form a BIBD($v=mn, b=nt, r, k=rm/t, \lambda$) ($\rightarrow 66$ Combinatorics).

As a generalization of Fisher's inequality, for an unequal-replicated block design with unequal block sizes, $b \geq v - \delta$ holds, where δ is

the multiplicity of the maximum eigenvalue 1 of the matrix $D_r^{-1/2} C D_r^{-1/2}$. The equality sign holds if and only if the projection operator corresponding to zero eigenvalue of $D_r^{-1/2} C D_r^{-1/2}$ is a zero matrix.

F. Estimation in a Mixed Model

Consider a block design (4), where every block has the same size k and every treatment has the same number r of replications. Let ξ be a fixed effect and η a random effect denoted by \mathbf{H} , and assume that \mathbf{W} satisfies conditions (a) and (b) of Section A and that \mathbf{H} satisfies conditions (d), (e), and (f) (where the \mathcal{E}_j are replaced by the coordinates H_j of \mathbf{H}). If $E(H_j) = \gamma$, $j=1, \dots, b$, then, changing the notation $\mathbf{H} - E_{v1}\gamma$ to \mathbf{H} , we can rewrite (4) as

$$\mathbf{Y} = \Gamma\gamma + \Phi\xi + \Psi\mathbf{H} + \mathbf{W}, \quad (8)$$

with $E(\mathbf{H}) = \mathbf{0}$. The normal equation that gives the least square estimate of ξ is

$$(C + \sigma^2(\sigma^2 + k\sigma_1^2)^{-1}C_1)\hat{\xi} \\ = \mathbf{Q} + \sigma^2(\sigma^2 + k\sigma_1^2)^{-1}\mathbf{Q}_1, \quad (9)$$

where C and \mathbf{Q} are the same as in (6) and $C_1 = ND_k^{-1}N' - rv^{-1}E_{vv}$, $\mathbf{Q}_1 = (ND_k^{-1}\Psi' - v^{-1}E_{v1}\Gamma')\mathbf{Y}$. Equation (9) cannot be solved unless the ratio $\sigma^2:\sigma_1^2$ is given. When $\sigma^2:\sigma_1^2$ is not known, substituting in $\sigma^2 + k\sigma_1^2$ its \dagger unbiased estimate given by analysis of variance, one obtains a solution of (9) which tends to a \dagger consistent estimate of ξ as the number of blocks tends to infinity.

G. Estimation in a Random-Effects Model

Let ξ (denoted by Ξ) and \mathbf{H} in the model (8) be random effects. Suppose that Ξ , \mathbf{H} , and \mathbf{W} are mutually independent, and that the distributions of Ξ , \mathbf{H} , and \mathbf{W} are $N(\mathbf{0}, \sigma_\xi^2 I_v)$, $N(\mathbf{0}, \sigma_1^2 I_b)$, and $N(\mathbf{0}, \sigma^2 I_n)$, respectively. The distribution of \mathbf{Y} in (8) contains four parameters $\gamma, \sigma^2, \sigma_1^2, \sigma_\xi^2$. When $k_j < v$, the \dagger minimal sufficient statistic is generally incomplete, and therefore the optimal estimate of $\gamma, \sigma^2, \sigma_1^2$, and σ_ξ^2 cannot be determined. As an example, the minimal sufficient statistic for the random-effects model of a BIBD(v, b, r, k, λ) is

$$(\sum_i Y_i, \mathbf{Y}' P_{21} \mathbf{Y}, \mathbf{Y}' P_{22} \mathbf{Y}, \mathbf{Y}' P_3 \mathbf{Y}, \mathbf{Y}' P_4 \mathbf{Y}, \mathbf{Y}' P_5 \mathbf{Y}),$$

where

$$P_{21} = k^{-1}(r-\lambda)^{-1}BTB - kr(r-\lambda)^{-1}n^{-1}E_{nn},$$

$$P_{22} = k((k-1)r + \lambda)^{-1}r^{-1}$$

$$\times (T - k^{-1}BT)(T - k^{-1}TB),$$

$$P_3 = k^{-1}B - k^{-1}(r-\lambda)^{-1}BTB$$

$$+ v\lambda(r-\lambda)^{-1}n^{-1}E_{nn},$$

$$P_4 = I_n - k^{-1}B - P_{22}, \quad P_5 = r^{-1}T - n^{-1}E_{nn},$$

with $T = \Phi\Phi'$ and $B = \Psi\Psi'$. In this case, $E(Y'P_4Y) = (n - v - b + 1)\sigma^2$, $E(Y'P_3Y) = (b - v)(\sigma^2 + k\sigma_1^2)$. From these equations unbiased estimates of σ^2 and σ_1^2 can be derived, but their optimality is not guaranteed (\rightarrow 396 Statistic).

H. Factorial Experiments

Suppose that there are h factors F_1, \dots, F_h which affect Y , and each factor F_i has s_i levels ($i = 1, \dots, h$). It is assumed that $v = s_1 \times s_2 \times \dots \times s_h$ treatment combinations are derived by all the combinations of the levels of h factors, and that v treatment effects are represented by the sum of subeffects called **main effects** and **interactions**, or **factorial effects** covering both. Such an experiment is called a **factorial experiment** or **factorial design**; it allows us to examine the effects of two or more factors, each factor being applied at two or more levels, by testing all possible treatment combinations formed from the factors under study. Specifically, suppose that we have the case $h = 2$, called a **two-way layout**. Let main effects be denoted by $\xi^1 = (\xi_1^1, \dots, \xi_{s_1}^1)$, $\xi^2 = (\xi_1^2, \dots, \xi_{s_2}^2)$, and interaction by $\xi^{12} = (\xi_{11}^{12}, \xi_{12}^{12}, \dots, \xi_{s_1 s_2}^{12})$, where $\sum_i \xi_i^1 = 0$, $\sum_j \xi_j^2 = 0$, $\sum_i \xi_i^{12} = \sum_j \xi_j^{12} = 0$. When there is no restriction on the number of observations, each of $v = s_1 \times s_2$ treatment combinations is replicated t times. Components Y_{ijk} of the observation vector Y are represented by a linear model

$$Y_{ijk} = \gamma + \xi_i^1 + \xi_j^2 + \xi_{ij}^{12} + W_{ijk},$$
$$i = 1, \dots, s_1; \quad j = 1, \dots, s_2; \quad k = 1, \dots, t,$$

or, in vector notation,

$$Y = \Gamma\gamma + X_1\xi^1 + X_2\xi^2 + X_{12}\xi^{12} + W.$$

The analysis of variance in this case is given by

$$Y'Y = \sum_i Y'P_iY,$$

where P_i is the projection operator matrix on the subspace derived by a direct sum decomposition

$$R^n = L(\Gamma) + L_\Gamma^\perp(X_1) + L_\Gamma^\perp(X_2) + L_{X_1, X_2}^\perp(X_{12}) + L_{X_{12}}^\perp,$$

where $n = s_1 s_2 t$. Denoting by $\bar{Y}_{ij\cdot}$ the arithmetic mean of Y_{ijk} over the subscript k and using similar notation $\bar{Y}_{i\cdot\cdot}$, $\bar{Y}_{\cdot j\cdot}$, and $\bar{Y}_{\cdot\cdot\cdot}$, we have $Y'P_1Y = n\bar{Y}_{\cdot\cdot\cdot}^2$, $Y'P_2Y = s_2 t \sum (\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot\cdot})^2$, $Y'P_3Y = s_1 t \sum (\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdot\cdot\cdot})^2$, $Y'P_4Y = t \sum (\bar{Y}_{ij\cdot} + \bar{Y}_{\cdot\cdot\cdot} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot j\cdot})^2$, $Y'P_5Y = \sum (Y_{ijk} - Y_{ij\cdot})^2$. The analysis of variance table is given in Table 1. For $h \geq 3$ similar models can also be considered for dealing with interactions up to h -factor. The factorial experiment is said to be **symmetric** if $s_1 = s_2 = \dots = s_h = s$ and is called the s^h **factorial experiment**. Otherwise it is

Table 1 Analysis of Variance

Source	Sum of Squares	Degrees of Freedom
F_1 main effects	$Y'P_2Y$	$s_1 - 1$
F_2 main effects	$Y'P_3Y$	$s_2 - 1$
$F_1 F_2$ interactions	$Y'P_4Y$	$(s_1 - 1)(s_2 - 1)$
Error	$Y'P_5Y$	$s_1 s_2 (t - 1)$

said to be **asymmetric** and is called the $s_1 \times s_2 \times \dots \times s_h$ **factorial experiment**.

I. Fractional Factorial Designs

Factorial designs require at least an experiment for all the combinations of levels of factors under consideration. When there are a large number of treatment combinations resulting from a large number of factors to be tested, it will often be beyond the resources of the investigator to experiment with all of them. For such cases Finney (*Ann. Eugen.*, 12 (1945)) proposed a method in which only a fraction of the treatment combinations are experimented with. A design of this type is called a **fractional factorial design**. One reason for the usefulness of fractional factorial designs in preference to factorial designs is that they involve a smaller number of treatment combinations, since, in most scientific experiments, it is usually found that a large number of the higher-order interactions are negligible. The crucial part of the specification of fractional factorial designs is the suitable choice of the defining or identity relationship. Equating the nonestimable factorial effects for the selected fraction of treatment combinations with I gives the **identity relation**. Each factorial effect is not estimable after selecting a fraction of treatment combinations, and any contrast of the selected treatment combinations represents more than one factorial effect. All factorial effects represented by the same treatment combinations are called **aliases**. In aliases, by assuming that other interactions are negligible when compared to the one of interest, estimation can be made by means of the corresponding contrast of the selected treatment combinations. When all factorial effects of order higher than l are assumed to be zero, a fractional factorial design is said to be of **resolution $2l + 1$** if it satisfies the condition that under the usual model all factorial effects up to order l are estimable, whereas a fractional factorial design is said to be of **resolution $2l$** if it satisfies the condition that under the usual model all factorial effects up to order $l - 1$ are estimable.

In the beginning, the theory was developed

for an **orthogonal fractional factorial design** in which the estimates of various effects of interest are all uncorrelated. However, these are available only for special values of the number n of treatment combinations and are in general uneconomical in the sense that they require a large value of n in comparison with the number of unknown effects. As a generalization of orthogonal fractional factorial designs, I. M. Chakravarti (*Sankhyā*, 17 (1956)) introduced the concept of **balanced fractional factorial designs**. In these designs, the estimates of effects are not always uncorrelated. However, the \dagger covariance matrix of the estimates is invariant under any permutation of factors. Balanced fractional factorial designs are flexible in the number of treatment combinations, with the result that more experimental situations can be handled. These two kinds of fractional factorial designs can be constructed by using orthogonal arrays and balanced arrays, defined in Section L. All alias relations can be derived from the identity relation in orthogonal designs.

In the model (1) for a design, the variance of the estimates of estimable linear functions of ξ depends on the matrix $X'X$ which is called the **information matrix**. For eigenvalues of $X'X$, the D -, E -, and A -optimalities of fractional factorial designs can be defined similarly to (I), (II), and (III) given in Section E. An orthogonal fractional 2^h factorial design is D -, E - and A -optimal. A fractional factorial design is said to be **saturated** if the number of treatment combinations is equal to that of parameters in the model to be estimated. However, since saturated designs do not provide an estimate of experimental error, their use should, in general, be confined to those situations in which a prior estimate of experimental error is available. The existence of a symmetric BIBD($4t-1, 4t-1, 2t-1, 2t-1, t-1$) implies that of an orthogonal saturated fractional 2^{4t-1} factorial design of resolution III.

J. Application of Algebras

In the theory of design of experiments, the ideas of association algebra and relationship algebra play an important role. Let A_i be a $v \times v$ symmetric matrix with entries 0 or 1 ($i = 0, 1, \dots, m$). If a set $\{A_i | i = 0, 1, \dots, m\}$ satisfies the conditions

$$A_0 = I_v, \quad \sum_{i=0}^m A_i = E_{vv},$$

and there is a nonnegative integer p_{jk}^i for every i, j, k such that

$$A_j A_k = \sum_{i=0}^m p_{jk}^i A_i,$$

then the A_i are called **association matrices**. In this case, there exists a natural number n_i for every i such that $A_i E_{vv} = n_i E_{vv}$ holds. If an (α, β) component $a_{\alpha\beta}^i$ of A_i is unity, then the treatments α and β are said to be the i th associates. The \dagger algebra \mathcal{A} over the real number field generated by the matrices A_0, A_1, \dots, A_m is called the **association algebra**. \mathcal{A} is commutative and \dagger completely reducible. $A_k \rightarrow \mathcal{P}_k = (p_{jk}^i)$ is the \dagger regular representation of \mathcal{A} . There is a nonsingular matrix $U = (u_{ij})$ that transforms all \mathcal{P}_i into diagonal matrices simultaneously:

$$U \mathcal{P}_i U^{-1} = \text{diag}(z_{0i}, \dots, z_{mi}), \quad i = 0, 1, \dots, m.$$

$$A_i^\# = \left(\sum_{f=0}^m u_{if} z_{if} \right)^{-1} \sum_{j=0}^m u_{ij} A_j, \quad i = 0, 1, \dots, m,$$

are mutually orthogonal idempotent elements of \mathcal{A} . For example, consider the case where $A_0 = I_{s_1} \otimes I_{s_2}$, $A_1 = (E_{s_1 s_1} - I_{s_1}) \otimes I_{s_2}$, $A_2 = I_{s_1} \otimes (E_{s_2 s_2} - I_{s_2})$, $A_3 = (E_{s_1 s_1} - I_{s_1}) \otimes (E_{s_2 s_2} - I_{s_2})$ with $m = 3$, where \otimes means the \dagger Kronecker product. The algebra \mathcal{A} generated by A_0, A_1, A_2, A_3 is called an F_2 type association algebra. The association matrices A_0, A_1, A_2 , and A_3 correspond to the relationships between the treatments in a two-way layout. The mutually orthogonal idempotent elements in this case are $A_0^\# = s_1^{-1} E_{s_1 s_1} \otimes s_2^{-1} E_{s_2 s_2}$, $A_1^\# = (I_{s_1} - s_1^{-1} E_{s_1 s_1}) \otimes s_2^{-1} E_{s_2 s_2}$, $A_2^\# = s_1^{-1} E_{s_1 s_1} \otimes (I_{s_2} - s_2^{-1} E_{s_2 s_2})$, and $A_3^\# = (I_{s_1} - s_1^{-1} E_{s_1 s_1}) \otimes (I_{s_2} - s_2^{-1} E_{s_2 s_2})$. For the factorial experiment with h factors ($h \geq 3$), association matrices can be constructed in a similar way. If $h = 3$, the number m of association matrices is 7. Many types of association schemes are known, and some of them for $m = 2$ are classified as group divisible, triangular, Latin square, cyclic type, and so on. For a group divisible type of $v = s_1 s_2$, $A_0 = I_v$, $A_1 = I_{s_1} \otimes E_{s_2 s_2} - A_0$, $A_2 = E_{vv} - A_0 - A_1$; $A_0^\# = v^{-1} E_{vv}$, $A_1^\# = v^{-1} \{(s_1 - 1)(A_0 + A_1) - A_2\}$, $A_2^\# = s_2^{-1} \{(s_2 - 1)A_0 - A_1\}$.

An experimental design consists of a set of n experimental units called plots. Define a relationship R between the plots as a set of ordered pairs (i, j) of plots. A relationship R among a set of n plots can be expressed as a symmetric $n \times n$ matrix (r_{ij}) of 0's and 1's:

$$r_{ij} = \begin{cases} 1 & \text{if the } i\text{th plot is related to the } j\text{th} \\ & \text{plot by the relationship } R, \\ 0 & \text{otherwise,} \end{cases}$$

and this matrix is also denoted by R . If there are k types of relationships R_1, \dots, R_k among n plots, the algebra \mathcal{R} over the real number field generated by the matrices R_1, \dots, R_k is called the **relationship algebra** of the design. \mathcal{R} is a \dagger semisimple algebra.

Example (1). The relationship algebra \mathcal{R} of the factorial experiment with $h = 2$ and t replications is generated by the following matrices

over the real number field: $B_1 = A_0 \otimes I_t$, $B_2 = A_0 \otimes (E_{tt} - I_t)$, $B_3 = A_1 \otimes E_{tt}$, $B_4 = A_2 \otimes E_{tt}$, $B_5 = A_3 \otimes E_{tt}$, where A_i , $i = 0, 1, 2, 3$, are the association matrices of F_2 type. The mutually orthogonal idempotents that correspond to the two-sided ideal decomposition of \mathcal{R} are $B_1^\# = A_0^\# \otimes t^{-1}E_{tt}$, $B_2^\# = A_1^\# \otimes t^{-1}E_{tt}$, $B_3^\# = A_2^\# \otimes t^{-1}E_{tt}$, $B_4^\# = A_3^\# \otimes t^{-1}E_{tt}$, and $B_5^\# = A_0 \otimes (I_t - t^{-1}E_{tt})$. These are the same as the projection P_i in the two-way layout design given in Section H.

Example (2). Consider a block design with v treatments, each having the same number r of replications, and with b blocks, each having the same size k ($< v$). Suppose that association matrices A_i , $i = 0, 1, \dots, m$, are given, by which the associations among the treatments are defined. Let $\lambda_{\alpha\beta}^i$ be the number of blocks to which the i th associate treatments α and β are applied. The design is called the **partially balanced incomplete block design (PBIBD)** if $\lambda_{\alpha\beta}^i = \lambda_i \geq 0$ independently of α and β . When $m = 1$, the design is a BIBD. Let the observation vector be represented by (4). The relationship algebra \mathcal{R} of a PBIBD is generated by $n \times n$ matrices I_n , E_{nn} , $B = \Psi\Psi'$, $T_i = \Phi A_i \Phi'$, $i = 1, 2, \dots, m$, where $NN' = \sum_{i=0}^m \lambda_i A_i = \sum_{i=0}^m \rho_i A_i^\#$, $\rho_i = \sum_{j=0}^m \lambda_j z_{ij}$ ($0 \leq \rho_i \leq rk$; $i = 0, 1, \dots, m$), and $\rho_0 = rk = \sum_{i=0}^m n_i \lambda_i$. Write $T_i^\# = \Phi A_i^\# \Phi'$. According as $\rho_i = rk$, $0 < \rho_i < rk$, or $\rho_i = 0$, $L(T_i^\#)$ is said to be **confounded** with the blocks, **partially confounded** with the blocks, or **orthogonal** to the blocks. \mathcal{R} is noncommutative, completely reducible, and isomorphic to the algebra of matrices of the type shown in Fig. 1.

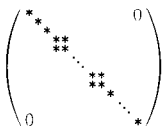


Fig. 1

The analysis of variance of a PBIBD is given by a decomposition of I_n into mutually orthogonal idempotent elements of \mathcal{R} . For a PBIBD, the matrix C in (6) is of the form $C = \sum_i \tau_i A_i^\#$, where $\tau_i = r - k^{-1} \rho_i$. For a connected PBIBD, all $A_i^\# \xi$, $i = 1, 2, \dots, m$, are estimable. If there exists a group divisible PBIBD ($m = 2$) with $\lambda_2 = \lambda_1 + 1$, then it is E - and A -optimal.

K. Design for Two-Way Elimination of Heterogeneity

Consider a design with v treatments in a $u \times w$ rectangular block. The row effect and the column effect of this block are denoted by η and ν , respectively. Thus the observation vector \mathbf{Y} is of the form

$$\mathbf{Y} = \Gamma\gamma + \Phi\xi + \Psi\eta + \Pi\nu + \mathbf{W}, \quad (10)$$

where the definitions of Γ , Φ , Ψ and Π are similar to those for block designs. Set $L = \Phi'\Pi$, $M = \Phi'\Psi$, $D_r = \Phi'\Phi$, $F = D_r - w^{-1}LL' - u^{-1}MM' + (uw)^{-1}LE_{uu}L'$. The matrix F plays a role similar to that of the matrix C in (6). When the rank of F equals $v - 1$, the design is said to be **connected**. If

$$F = \tau(I_v - v^{-1}E_{vv}), \quad (11)$$

then the design satisfies the optimum criteria (I), (II), and (III) given in Section E. When $u = w = v$ and (11) holds, the design is called the **Latin square**. When $u = v$ and (11) holds, the design is called the **Youden square**. When $u > v$ and (11) holds, the design is called the **Shrikhande square**. In a **Youden square design** treatments are grouped into replications in two different ways, i.e., rows and columns, where rows constitute a BIBD, whereas columns are complete blocks. The existence of a Youden square design is equivalent to that of a symmetric BIBD.

If the associations among v treatments are defined in terms of association matrices, the partially balanced design for **two-way elimination of heterogeneity** can be defined in a way similar to the PBIBD. In this case, the equation $F = \sum_i \tau_i A_i^\#$ holds, and if $\mathcal{A} = \{I_v, E_{vv} - I_v\}$, then (11) holds, hence the optimum criteria are fulfilled. The definition of the relationship algebra \mathcal{R} of a partially balanced design for two-way elimination of heterogeneity is similar to that used for the PBIBD. \mathcal{R} is isomorphic to the algebra of matrices shown in Fig. 2.

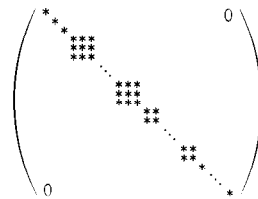


Fig. 2

The analysis of variance of this design is given by a decomposition of I_n into mutually orthogonal idempotent elements of \mathcal{R} .

L. Balanced Array and Orthogonal Array

Suppose that T is an $n \times h$ matrix with entries from a set A of s (≥ 2) distinct elements. Consider the $s^t \times t$ matrices $X = (x_1, x_2, \dots, x_t)$ that can be formed by giving different values to the x_i ($\in A$), $i = 1, 2, \dots, t$. Suppose that associated with each $1 \times t$ matrix X there is a non-negative integer $\lambda(x_1, x_2, \dots, x_t)$ which is invariant under any permutation of a given set $\{x_1, x_2, \dots, x_t\}$. If, for every t -columned submatrix of T , the $s^t \times t$ matrices X occur as

rows $\lambda(x_1, x_2, \dots, x_t)$ times, then the matrix T is called the **balanced array (BA)** of size n , h constraints, s levels, and strength t having the index set $\{\lambda(x_1, \dots, x_t)\}$. Such an array is denoted by $BA(n, h, s, t)$; n is also called the **number of treatment combinations**. In particular, when $\lambda(x_1, \dots, x_t) = \lambda$ for every x_1, \dots, x_t , such an array is called an **orthogonal array (OA)** of size n , h constraints, s levels, strength t , and index λ , which is denoted by $OA(n, h, s, t)$. We have $n = \lambda s^t$. Balanced arrays have the advantage that they can be constructed with fewer treatment combinations than the orthogonal arrays for given h, s, t parameters. The transpose of the incidence matrix of a BIBD(v, b, r, k, λ) is, for $A = \{0, 1\}$, a BA($b, v, 2, 2$) with $\lambda(0, 0) = b - 2r + \lambda$, $\lambda(1, 0) = \lambda(0, 1) = r - \lambda$, and $\lambda(1, 1) = \lambda$. There is a close relation between the existence of an OA and that of a BIBD. For example, the existence of an $OA(s^2, h, s, 2)$ is equivalent to that of $h - 2$ mutually orthogonal Latin squares of order s that is for $h = s + 1$ equivalent to the existence of a BIBD($s^2, s(s + 1), s + 1, s, 1$). The existence of an $OA(4t, 4t - 1, 2, 2)$ is equivalent to that of a symmetric BIBD($4t - 1, 4t - 1, 2t - 1, 2t - 1, t - 1$) which is also equivalent to the existence of an $OA(8t, 4t, 2, 3)$.

Balanced arrays and orthogonal arrays play a vital role in the construction of symmetric and asymmetric confounded factorial experiments and fractional factorial designs. In an $OA(n, h, s, t)$, regarding columns and entries as h factors and levels of a factor of the column, respectively, each row corresponds to a treatment combination. In this case, this orthogonal array gives an orthogonal fractional s^t factorial design of resolution $t + 1$ with n treatment combinations. A necessary and sufficient condition for a fractional s^h factorial design of resolution $2l + 1$ to be orthogonal (resp. balanced) is that the design be an orthogonal (resp. balanced) array of strength $2l$, provided that the information matrix of the design is nonsingular.

M. Response Surface

If all the factors represent quantitative variables, such as time, temperature, amount of ingredients, etc., it is natural to think of the yield or response Y of the experimental results as a continuous function of the levels of these factors. We can write a functional relationship

$$Y_\alpha = f(x_{1\alpha}, x_{2\alpha}, \dots, x_{h\alpha}) + W_\alpha, \quad \alpha = 1, 2, \dots, n,$$

where Y_α represents the α th observation in the designed experiment and $x_{i\alpha}$ represents the level of the i th factor in the α th observation. The function f or the surface defined in $(h + 1)$ -

dimensional Euclidean space by f is called the **response surface**. The residual W_α measures the experimental error of the α th observation. A knowledge of the function f gives a complete summary of experimental results and also enables us to predict the response for values of the $x_{i\alpha}$ that were not tested in the experiment, or to determine such combinations of the values of variables $x_{i\alpha}$ that give the maximum (or minimum) value of f . When the mathematical form of f is not known, the function can sometimes be approximated satisfactorily within the experimental region, which is relatively small, by flexible graduating functions, such as polynomials of degrees 1 and 2 in the variables $x_{i\alpha}$:

$$Y_\alpha = \beta_0 + \sum_{i=1}^h \beta_i x_{i\alpha} + W_\alpha,$$

$$Y_\alpha = \beta_0 + \sum_{i=1}^h \beta_i x_{i\alpha} + \sum_{i=1}^h \beta_{ii} x_{i\alpha}^2 + \sum_{i < j} \beta_{ij} x_{i\alpha} x_{j\alpha} + W_\alpha.$$

The coefficients β_0, β_1, \dots are parameters to be estimated from the data. Thus, (i) when the form of the true f is assumed known, the object is to estimate the parameters; (ii) when the form of the true f is unknown, the object is to approximate the f by some graduating function. Designs appropriate for (i) and (ii) are called **designs for estimating parameters** and **designs for exploring a response surface**, respectively. Some experimental designs that have been developed for fitting polynomials of the first and second degrees are called **first-order designs** and **second-order designs**, respectively. The problem here is to increase the precision of fitting response surfaces by appropriately choosing n points in a given experimental region S of variable x . That is, how do we find the experimental region that interests us in the sense of finding optimum conditions and, having found it, how do we design experiments to map f over the region? As optimum criteria of allocations, corresponding to (I) and (II) given in Section E, consider the following: (IV) the generalized variance of the estimates of coefficients is minimal; (V) the supremum of the variance of estimates of the expectation $E(Y)$ of response Y is minimal for all variables $x \in S$. For example, when an observation Y is given by a polynomial regression of a variable x as $Y_\alpha = \beta_0 + \beta_1 x_\alpha + \beta_2 x_\alpha^2 + \dots + \beta_h x_\alpha^h + W_\alpha$, if for the Legendre polynomial $P_h(x)$ we allocate an experiment replicated at $x = \pm 1$ and $h - 1$ roots of $P_h(x) = 0$ the same number of times, then it is an optimum design in the sense of (IV) and (V).

As a design for fitting first-order response planes, we can use a 2^h factorial design (or its fraction) of resolution III, since it is sufficient to consider only linear effects for each of h

factors. In comparing first-order designs, we take as an optimum criterion the minimum of the average variance of the estimates of the coefficients, corresponding to (III) of Section E. For a design for fitting second-order response surfaces, we need at least three levels of each factor for estimating all effects up to quadratic ones; then we can use a 3^h factorial design (or its fraction) of resolution V for estimating the linear \times linear component of interactions of two factors. In this case, treatment combinations are too large in number, so we can adopt **central composite designs**, which are constructed by adding further treatment combinations to those obtained from a 2^h factorial design (or its fraction). If the coded levels of each factor are -1 and $+1$ in the 2^h factorial design, the $(2h+1)$ additional combinations are $(0, 0, \dots, 0)$, $(\pm d, 0, \dots, 0)$, $(0, \pm d, 0, \dots, 0)$, \dots , $(0, \dots, 0, \pm d)$. The total number of treatment combinations to be tested is $2^h + 2h + 1$ ($< 3^h$ for $h \geq 3$). The value of d can be chosen to make the coefficients in the quadratic polynomials as orthogonal as possible to each other or to minimize the bias that is created if the true form of the response surface is not quadratic or to give the design with the property of being rotatable. If \hat{Y}_0 is the estimated response at $(x_{10}, x_{20}, \dots, x_{h0}) \in S$, the response surface design is said to be **rotatable** if and only if the variance of \hat{Y}_0 is a function only of the distance $\rho = (x_{10}^2 + x_{20}^2 + \dots + x_{h0}^2)^{1/2}$ of $(x_{10}, x_{20}, \dots, x_{h0})$ from the origin, so that the variance contours in the experimental region of the variables are circles, spheres, or hyperspheres centered at the origin. Now, when $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + 2\beta_{12} x_1 x_2 + W$ with $|x_i| \leq 1$ for $i = 1, 2$, an optimum experiment in the sense of (V) is given by allocating 0.0960 at the origin, 0.0802 at experimental points $(1, 0)$, $(-1, 0)$, $(0, 1)$, $(0, -1)$, and 0.1458 at experimental points $(1, 1)$, $(1, -1)$, $(-1, 1)$, $(-1, -1)$, respectively. However, this design is not rotatable. Instead of an optimum criterion (V), we can consider a region \tilde{S} which is not always equal to the original experimental region S and a criterion: (V') The supremum of the variance of estimates of the expectation $E(Y)$ of response Y is minimal for all variables $x \in \tilde{S}$. Within $x_1^2 + x_2^2 \leq c^2$, an optimum allocation depends on the value of c , and it is shown that a rotatable design is obtained only if c is in a certain range.

There are many topics in the theory of experimental design besides the ones mentioned in this article (\rightarrow [4] for multiple comparison, [9] for confounding designs that are factorial experiments in which the block size is reduced and in which for each block a fractional of all the treatment combinations are tested, [10] for split-plot designs in which certain main effects

are confounded with blocks, [15] for weighing designs).

References

- [1] "Experimental design" (in Japanese), Gen-dai tōkeigaku daiziten (Encyclopedia of modern statistics), Tōyōkeizai simpōsyū, 1962.
- [2] R. A. Fisher, The design of experiments, Oliver & Boyd, seventh edition, 1960.
- [3] P. W. M. John, Statistical design and analysis of experiments, Macmillan, 1971.
- [4] H. Scheffé, The analysis of variance, Wiley, 1959.
- [5] J. Kiefer, Optimum experimental designs, J. Roy. Statist. Soc., ser. B, 21 (1959), 273–319.
- [6] V. V. Fedorov, Theory of optimal experiments, Academic Press, 1972.
- [7] D. Silvey, Optimal design, Chapman, 1980.
- [8] D. Raghavarao, Constructions and combinatorial problems in design of experiments, Wiley, 1971.
- [9] O. Kempthorne, The design and analysis of experiments, Wiley, 1952.
- [10] W. G. Cochran and G. M. Cox, Experimental designs, Wiley, 1950.
- [11] A. T. James, The relationship algebra of an experimental design, Ann. Math. Statist., 28 (1957), 993–1002.
- [12] G. E. P. Box and N. R. Draper, Evolutionary operation, Wiley, 1968.
- [13] M. Kendall and A. Stuart, The advanced theory of statistics, vol. 3, Design and analysis, and time-series, Griffin, third edition, 1976.
- [14] G. E. P. Box and K. B. Wilson, On the experimental attainment of optimum conditions, J. Roy. Statist. Soc., ser. B, 13 (1952), 1–45.
- [15] K. S. Banerjee, Weighing designs for chemistry, medicine, economics, operation research, statistics, Dekker, 1975.

103 (III.3) Determinants

A. Definition

The **determinant** of an $n \times n$ matrix $A = (a_{ik})$ in a \dagger commutative ring R is defined to be the following element of R :

$$\sum (\text{sgn } P) a_{1p_1} a_{2p_2} \dots a_{np_n},$$

where

$$P = \begin{pmatrix} 1 & 2 & \dots & n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

is a permutation of the numbers $1, 2, \dots, n$,

$\text{sgn } P$ denotes the **sign** of the permutation P (that is, $\text{sgn } P = 1$ if P is an even permutation and $\text{sgn } P = -1$ if P is an odd permutation), and the summation extends over all $n!$ permutations of $1, 2, \dots, n$. The determinant of A is denoted by

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

It is written $|a_{ik}|$ or $|A|$ and is also denoted by $\det A$. Usually we suppose that R is the field \mathbf{R} of real numbers or the field \mathbf{C} of complex numbers, but the following theorems are also valid for cases in which R is any commutative ring, unless otherwise stated.

B. Relation to Exterior Algebras

Consider an r -exterior algebra (r -Grassmann algebra) of a linear space (r -free module) of dimension n over R with a basis (e_1, e_2, \dots, e_n) . Set

$$e'_i = a_{i1}e_1 + a_{i2}e_2 + \cdots + a_{in}e_n,$$

where $a_{ij} \in R$. Then we have $e'_1 \wedge e'_2 \wedge \cdots \wedge e'_n = |a_{ik}| e_1 \wedge e_2 \wedge \cdots \wedge e_n$. Conversely, we can define the determinant $|a_{ik}|$ by this relation. The properties of determinants can be easily deduced from those of exterior algebras.

C. Fundamental Properties of Determinants

(1) The determinant of the r -transposed matrix tA of a matrix A is equal to the determinant of A . Hence the theorems stated for rows are also valid for columns.

(2) If the elements of one row (column) of a matrix are multiplied by a factor c , the determinant of the matrix is also multiplied by c . If the elements of one row (column) of a matrix are zero, its determinant is equal to zero.

(3) If from a matrix $A = (a_{ik})$, we obtain two matrices A' and A'' by replacing one row, for instance the i th row, by a'_{i1}, \dots, a'_{in} and by $a_{i1} + a'_{i1}, \dots, a_{in} + a'_{in}$, respectively, then $|A''| = |A| + |A'|$. This relation is equally valid for a column.

(4) If we obtain A_Q by a permutation Q on the rows of a matrix A , then $|A_Q| = (\text{sgn } Q)|A|$. In particular, if two rows (columns) of a matrix are interchanged, then the determinant changes sign.

(5) The determinant of a matrix is zero if two rows (or columns) are identical.

(6) The determinant of a matrix is not changed if the elements of any row (column), each multiplied by the same factor, are added

to the corresponding elements of another row (column).

(7) Suppose that R has unity element. Let x_{ik} ($i, k = 1, \dots, n$) be n^2 variables in R , and denote a function (having its values in R) of these variables by $\varphi(X)$, $X = (x_{ik})$. Assume that $\varphi(X)$ has the following properties: (i) if the elements of one row of X are multiplied by a factor λ , the value of φ is also multiplied by λ ; (ii) if we obtain two matrices X' and X'' by replacing one row of X , for instance the i th row, by x'_{i1}, \dots, x'_{in} and by $x_{i1} + x'_{i1}, \dots, x_{in} + x'_{in}$, respectively, then $\varphi(X'') = \varphi(X) + \varphi(X')$; and (iii) if two rows of X are equal, $\varphi(X) = 0$. Then $\varphi(X) = c|X|$ for some constant c (in R).

(8) Suppose now that R is a field K . Assume that a function $\varphi(X)$ (in K) has the following properties: (i) if the elements of one row in X are multiplied by λ , the value of φ is also multiplied by λ ; and (ii) the value of φ is not changed if the elements of any row are added to the corresponding elements of another row. Then $\varphi(X) = c|X|$ for some constant c (in K).

D. The Laplace Expansion Theorem

Let $A = (a_{ik})$ be an $n \times n$ matrix. Take r -tuples (i_1, \dots, i_r) and (k_1, \dots, k_r) , where i_a and k_b belong to $\{1, \dots, n\}$ and $i_1 < \dots < i_r$, $k_1 < \dots < k_r$. Let (i_{r+1}, \dots, i_n) and (k_{r+1}, \dots, k_n) be $(n-r)$ -tuples such that $i_{r+1} < \dots < i_n$, $k_{r+1} < \dots < k_n$ and $\{i_1, \dots, i_r, i_{r+1}, \dots, i_n\} = \{k_1, \dots, k_r, k_{r+1}, \dots, k_n\} = \{1, \dots, n\}$. Let $a_{(i_1, \dots, i_r)(k_1, \dots, k_r)}$ be the determinant of an $r \times r$ matrix whose (p, q) -component is the (i_p, k_q) -component of A for each p and q . We call this determinant a **minor** of degree r of the matrix A . (The corresponding submatrix of A is sometimes also called a minor of A .) In particular, if $(i_1, \dots, i_r) = (k_1, \dots, k_r)$, then it is called a **principal minor**. Furthermore, we define the **cofactor** of the minor $a_{(i_1, \dots, i_r)(k_1, \dots, k_r)}$ of A to be

$$\tilde{a}_{(i_1, \dots, i_r)(k_1, \dots, k_r)} = (-1)^{\lambda + \mu} a_{(i_{r+1}, \dots, i_n)(k_{r+1}, \dots, k_n)},$$

where $\lambda = i_1 + \dots + i_r$ and $\mu = k_1 + \dots + k_r$. In the particular case $r = 1$, the cofactor of a_{ik} is $\tilde{a}_{ik} = (-1)^{i+k} \Delta_{ik}$, where Δ_{ik} is the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by eliminating its i th row and k th column. For simplicity, we abbreviate (i_1, \dots, i_r) , (k_1, \dots, k_r) , and (j_1, \dots, j_r) as (i) , (k) , and (j) , respectively. Then we have

$$\sum_{(j)} a_{(i)(j)} \tilde{a}_{(k)(j)} = \begin{cases} |A| & \text{if } (i) = (k), \\ 0 & \text{if } (i) \neq (k), \end{cases}$$

$$\sum_{(j)} a_{(j)(k)} \tilde{a}_{(j)(i)} = \begin{cases} |A| & \text{if } (i) = (k), \\ 0 & \text{if } (i) \neq (k), \end{cases}$$

where $\sum_{(j)}$ means that the sum is taken over all combinations (j) . This is called the **Laplace expansion theorem**. If a matrix A has the form

$$A = \begin{pmatrix} B & 0 \\ * & C \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} B & * \\ 0 & C \end{pmatrix},$$

and B and C are square matrices, then by this theorem we have $|A| = |B||C|$. If we number the combinations $(i) = (i_1, \dots, i_r)$ and $(k) = (k_1, \dots, k_r)$ appropriately (for instance, in lexicographical order) and regard the numbers assigned to them as row numbers and column numbers, respectively, to form a matrix $(a_{(i)(k)})$, then the Laplace theorem can be expressed as

$$(a_{(i)(k)}) (\tilde{a}_{(k)(i)}) = (\tilde{a}_{(k)(i)}) (a_{(i)(k)}) \\ = \begin{pmatrix} |A| & & 0 \\ & \ddots & \\ 0 & & |A| \end{pmatrix}.$$

In the particular case $r = 1$, we have

$$\sum_{j=1}^n a_{ij} \tilde{a}_{kj} = \begin{cases} |A| & \text{if } i = k, \\ 0 & \text{if } i \neq k, \end{cases}$$

$$\sum_{j=1}^n a_{jk} \tilde{a}_{ji} = \begin{cases} |A| & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases}$$

E. Product of Determinants

Let $A = (a_{ik})$ and $B = (b_{ik})$ be two $n \times n$ matrices. For the product $AB = C = (c_{ik})$, where $c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$ ($i, k = 1, \dots, n$), we have $|AB| = |A||B|$. The \dagger inverse matrix A^{-1} exists for an $n \times n$ matrix $A = (a_{ik})$ if and only if $|A| \neq 0$, and then $A^{-1} = (b_{ik})$ with elements $b_{ik} = \tilde{a}_{ki}/|A|$. Moreover, we have $|A^{-1}| = |A|^{-1}$. (In the case where the elements a_{ik} are in the commutative ring R with unity element, A^{-1} exists if and only if $|A|$ is a \dagger regular element of R .)

F. Theorems on Determinants

(1) Let \tilde{a}_{ik} be the cofactor of a_{ik} in the determinant of an $n \times n$ matrix $A = (a_{ik})$. Then the determinant $|\tilde{a}_{ik}|$ is equal to $|A|^{n-1}$. In general,

$$|a_{(i_1, \dots, i_r)(k_1, \dots, k_r)}| = |A|^{\binom{n-1}{r-1}}, \\ |\tilde{a}_{(i_1, \dots, i_r)(k_1, \dots, k_r)}| = |A|^{\binom{n-1}{r}}.$$

(2) The determinant of a submatrix of the matrix (\tilde{a}_{ik}) , composed of the i_1 th, \dots , i_r th rows and k_1 th, \dots , k_r th columns of (\tilde{a}_{ik}) is equal to

$$|A|^{r-1} \tilde{a}_{(i_1, \dots, i_r)(k_1, \dots, k_r)}.$$

(3) Let $\Delta \begin{pmatrix} i_1, \dots, i_r \\ k_1, \dots, k_r \end{pmatrix}$ be the determinant

of the $(n-r) \times (n-r)$ matrix obtained from an $n \times n$ matrix A by eliminating its i_1 th, \dots , i_r th rows and k_1 th, \dots , k_r th columns. Then

$$|A| \Delta \begin{pmatrix} i & j \\ k & l \end{pmatrix} = \Delta \begin{pmatrix} i \\ k \end{pmatrix} \Delta \begin{pmatrix} j \\ l \end{pmatrix} - \Delta \begin{pmatrix} i \\ l \end{pmatrix} \Delta \begin{pmatrix} j \\ k \end{pmatrix}, \\ i < j, \quad k < l.$$

(4) **Sylvester's theorem.** Let b_{ik} ($i, k = 1, \dots, n-r$) denote the minor $a_{(1, \dots, r, r+i)(1, \dots, r, r+k)}$ of an $n \times n$ matrix $A = (a_{ij})$. Then

$$|b_{ik}| = |A| \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix}^{n-r-1}.$$

(5) Let A be an $n \times m$ matrix and B an $m \times n$ matrix. Then AB is an $n \times n$ matrix. If $n > m$, then $|AB| = 0$. If $n \leq m$, let $(i) = (i_1, \dots, i_n)$ ($i_1 < \dots < i_n$) be a combination of $1, 2, \dots, m$, taken n at a time. Let $A_{(i)}$ be the $n \times n$ matrix composed of the i_1 th, \dots , i_n th columns of A , and $B_{(i)}$ the $n \times n$ matrix composed of the i_1 th, \dots , i_n th rows of B . Then $|AB| = \sum_{(i)} |A_{(i)}| |B_{(i)}|$, where the summation extends over all possible combinations (i) .

(6) **Determinant of a \dagger Kronecker product.** If A is an $m \times m$ matrix and B is an $n \times n$ matrix, then $|A \otimes B| = |A|^n |B|^m$.

(7) Let H be an $n \times n$ \dagger Hermitian matrix, and let H_k denote the matrix composed of its first k rows and columns. Then H is positive definite if and only if $|H_k| > 0$ for all $k = 1, \dots, n$.

G. Special Determinants

(1) A determinant of the form

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}$$

is called a **Vandermonde determinant**. It is equal to the \dagger simplest alternating function $\prod_{i>k} (x_i - x_k)$.

(2) A **cyclic determinant** is one of the form

$$\begin{vmatrix} x_0 & x_1 & x_2 & \dots & x_{n-1} \\ x_{n-1} & x_0 & x_1 & \dots & x_{n-2} \\ & & & \ddots & \\ x_1 & x_2 & x_3 & \dots & x_0 \end{vmatrix} \\ = \prod_{i=0}^{n-1} (x_0 + \zeta^i x_1 + \zeta^{2i} x_2 + \dots + \zeta^{(n-1)i} x_{n-1}),$$

where ζ is a \dagger primitive n th root of unity.

(3) Consider the vectors $\alpha_i = (a_{i1}, a_{i2}, \dots, a_{in})$ ($i = 1, 2, \dots, n$), and let (α_i, α_j) denote the \dagger inner product of α_i and α_j . Then the following determinant is called the **Gramian** of these vectors:

$$\begin{vmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \dots & (\alpha_1, \alpha_n) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \dots & (\alpha_2, \alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_n, \alpha_1) & (\alpha_n, \alpha_2) & \dots & (\alpha_n, \alpha_n) \end{vmatrix} \\ = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}^2.$$

(4) For an alternating matrix (namely, a square matrix X such that $X = -X^T$), we have the identity

$$\begin{vmatrix} 0 & x_{12} & x_{13} & \cdots & x_{1n} \\ -x_{12} & 0 & x_{23} & \cdots & x_{2n} \\ -x_{13} & -x_{23} & 0 & \cdots & x_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -x_{1n} & -x_{2n} & -x_{3n} & \cdots & 0 \end{vmatrix} = \begin{cases} P_n(\dots, x_{ij}, \dots)^2 & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd,} \end{cases}$$

where $P_n(\dots, x_{ij}, \dots)$ is a polynomial of variables x_{ij} , which (equipped with appropriate sign) is called the **Pfaffian** of these variables.

References

- [1] I. Satake, Linear algebra, Dekker, 1975.
- [2] G. W. H. Kowalewski, Einführung in die Determinantentheorie, de Gruyter, second revised edition, 1925 (Chelsea, 1948).
- [3] N. Bourbaki, Eléments de mathématique, Algèbre, ch. 3, Actualités Sci. Ind., 1044, Hermann, new edition, 1958.
- [4] F. R. Gantmakher, The theory of matrices, Chelsea, I, II, 1959. (Original in Russian, 1953.)
- [5] A. G. Kurosh, Lectures on general algebra, Chelsea, 1970. (Original in Russian, 1955.)
- [6] R. Godement, Cours d'algèbre, Hermann, 1963.
- [7] S. MacLane and G. Birkhoff, Algebra, Macmillan, 1967.

104 (XIII.15) Difference Equations

A. General Remarks

Let y be a function of a real variable x defined on an interval I and let Δx be a fixed quantity. When two points x and $x + \Delta x$ are in I , we define the **difference** $\Delta y(x)$ of y at x by $\Delta y(x) = y(x + \Delta x) - y(x)$ and the difference quotient by $\Delta y(x)/\Delta x$; Δx is called the difference of x . Without loss of generality, we can take $\Delta x = 1$, for otherwise there is a constant b such that $\Delta x' = 1$ for the new independent variable $bx = x'$. If $\Delta x = 1$, the **second difference** $\Delta^2 y(x) = \Delta(\Delta y(x))$ is given by

$$\begin{aligned} \Delta^2 y(x) &= \Delta y(x+1) - \Delta y(x) \\ &= y(x+2) - 2y(x+1) + y(x). \end{aligned}$$

Similarly, the **difference of the n th order** is defined by $\Delta^n y(x) = \Delta(\Delta^{n-1} y(x))$, and

$$\Delta^n y(x) = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} y(x+k).$$

Conversely, $y(x+n)$ is expressed by differences as

$$y(x+n) = \sum_{k=0}^n \binom{n}{k} \Delta^k y(x)$$

(\rightarrow 223 Interpolation).

B. Summation

Given a function $g(x)$ and Δx , a function $y(x)$ that satisfies $\Delta y(x)/\Delta x = g(x)$ is called a **sum** of $g(x)$. **Summation** of $g(x)$ is to find a sum of $g(x)$. Given a sum $y(x)$ of $g(x)$, an **indefinite sum** of $g(x)$, written as $Sg(x)\Delta x$, is given by $Sg(x)\Delta x = y(x) + c(x)$, where $c(x)$ is an arbitrary periodic function of period Δx . In many cases, $c(x)$, which corresponds to an arbitrary constant in an indefinite integral, is omitted. For example, a sum of $g(x) = nx^{n-1}$ for $\Delta x = 1$ is the n th-order \dagger Bernoulli polynomial $B_n(x)$ for $n \neq 0$; a sum of x^{-1} is $\psi(x)$, given by $\psi(x) = d \log \Gamma(x)/dx$ (\rightarrow 174 Gamma Function).

When the series $-\Delta x \sum_{k=0}^{\infty} g(x+k\Delta x)$ or $\Delta x \sum_{k=-1}^{\infty} g(x-k\Delta x)$ converges, both can be sums of $g(x)$. Since the requirement of convergence for these series was found to be too strict, the following requirement was given by N. E. Nörlund instead: Let x be a real variable and $g(x)$ be continuous for $x \geq b$. Define $\lambda(x)$ by $\lambda(x) = x^p (\log x)^q$ ($p \geq 1, q \geq 0$). Then if for a positive η

$$\begin{aligned} F(x, \Delta x, \eta) &= \int_a^{\infty} g(z) e^{-\eta \lambda(z)} dz \\ &\quad - \Delta x \sum_{k=0}^{\infty} g(x+k\Delta x) e^{-\eta \lambda(x+k\Delta x)} \end{aligned}$$

is convergent for $a > b$, F satisfies $\Delta F(x, \Delta x, \eta)/\Delta x = g(x) \exp(-\eta \lambda(x))$. Accordingly, if $F(x, \Delta x, \eta)$ approaches a limit $F(x)$ as $\eta \rightarrow 0$, $F(x)$ is a solution of $\Delta F(x)/\Delta x = g(x)$. $F(x)$ is called the **principal solution** of $\Delta F(x)/\Delta x = g(x)$.

C. Difference Equations

Let $\Delta x = 1$. An equation $F(x, y(x), \Delta y(x), \dots, \Delta^n y(x)) = 0$ in x and differences of an unknown function $y(x)$ is called a **difference equation**. If the substitution $y = \varphi(x)$ satisfies the equation for x in some interval, $\varphi(x)$ is a **solution** of the equation. Because of the relation between $y(x), y(x+1), \dots, y(x+n)$ and the differences of y at x , we can transform the given difference equation in the form $G(x, y(x), y(x+1), \dots, y(x+n)) = 0$. This form appears more often in applications and is called the **standard form** of a difference equation.

If the equation is linear in $y(x)$, $y(x+1)$, \dots , $y(x+n)$, namely, if it is given by

$$\sum_{i=0}^n p_i(x)y(x+i)=q(x),$$

the difference equation is said to be **linear**.

When $q(x) \equiv 0$, it is **homogeneous**; otherwise, it is **inhomogeneous** (or **nonhomogeneous**).

D. Linear Difference Equations

Assume that $p_0(x), \dots, p_n(x)$ are single-valued analytic functions without poles and common zeros in some domain. Consider the linear difference equation

$$\sum_{i=0}^n p_i(x)y(x+i)=0. \quad (1)$$

If $\varphi_1(x), \varphi_2(x), \dots, \varphi_m(x)$ are solutions of (1), then a linear combination $a_1(x)\varphi_1(x) + a_2(x)\varphi_2(x) + \dots + a_m(x)\varphi_m(x)$ with arbitrary periodic functions $a_1(x), a_2(x), \dots, a_m(x)$ of period 1 is also a solution of (1).

Let β_1, β_2, \dots be singular points of $p_1(x), p_2(x), \dots, p_n(x)$, $\alpha_1, \alpha_2, \dots$ be the zeros of $p_0(x)$, and $\gamma_1, \gamma_2, \dots$ be the zeros of $p_n(x+n)$. Then the set of **singular points** of the linear difference equation (1) is the set $\{\alpha_i, \beta_i, \gamma_i\}$.

A function $\varphi_m(x)$ is said to be **linearly dependent** on the functions $\varphi_1(x), \varphi_2(x), \dots, \varphi_{m-1}(x)$ with respect to the difference equation (1) if $\varphi_m(x) = a_1(x)\varphi_1(x) + a_2(x)\varphi_2(x) + \dots + a_{m-1}(x)\varphi_{m-1}(x)$, where $a_1(x), a_2(x), \dots, a_{m-1}(x)$ are functions of period 1, every one of which takes a nonzero finite value at least at one point not congruent (mod \mathbf{Z}) to any of the singular points, where \mathbf{Z} is the additive group of integers.

A set of m functions is called **linearly independent** if none of the functions is dependent on the other $m-1$ functions. When a set of n solutions of equation (1) is linearly independent, it is a **fundamental system** for (1). Any solution of (1) can be expressed as a linear combination of n solutions of a fundamental system.

The determinant

$$\begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \dots & \varphi_n(x) \\ \varphi_1(x+1) & \varphi_2(x+1) & \dots & \varphi_n(x+1) \\ \dots & \dots & \dots & \dots \\ \varphi_1(x+n-1) & \varphi_2(x+n-1) & \dots & \varphi_n(x+n-1) \end{vmatrix}$$

formed from n functions $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ is called **Casorati's determinant** and is denoted by $D(\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x))$. A necessary and sufficient condition for a given set of n functions to be independent is that Casorati's determinant be nonzero at every point except those which are congruent to singular points of (1). Casorati's determinant is used to deter-

mine whether a given set of solutions is fundamental.

Let $\psi(x)$ be a solution of a nonhomogeneous linear difference equation

$$P_x(y) = \sum_{i=0}^n p_i(x)y(x+i) = q(x). \quad (2)$$

If $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ are n linearly independent solutions of (1), then an arbitrary solution of (2) is given by

$$y = a_1(x)\varphi_1(x) + a_2(x)\varphi_2(x) + \dots + a_n(x)\varphi_n(x) + \psi(x),$$

where $a_1(x), \dots, a_n(x)$ are arbitrary periodic functions of period 1. Then the expression for y is called a **general solution** of (2). If we abbreviate Casorati's determinant of a fundamental system of solutions $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ of (1) by $D(x)$ and write $\mu_i(x)$ as the quotient of the cofactor of $\varphi_i(x+n)$ of $D(x+1)$ by $D(x+1)$, we have

$$\psi(x) = \sum_{i=1}^n \varphi_i(x) S q(z) \mu_i(z) \Delta z,$$

assuming that the summation S on the right-hand side is known. This is the analog of Lagrange's method of variation of constants in the theory of linear ordinary differential equations.

E. Linear Difference Equations with Constant Coefficients

If all the coefficients in

$$\sum_{i=0}^n p_i y(x+i) = 0, \quad p_0 \neq 0, \quad p_n \neq 0, \quad (3)$$

are constants, n linearly independent solutions are obtained easily. Indeed, if λ is a root of the algebraic equation $\sum_{i=0}^n p_i \lambda^i = 0$, λ^x is a solution of (3). This algebraic equation is called the **characteristic equation** of (3). If it has n distinct roots $\lambda_1, \lambda_2, \dots, \lambda_n$, then $\lambda_1^x, \lambda_2^x, \dots, \lambda_n^x$ are n linearly independent solutions. In general, if λ is an m -tuple root of the characteristic equation, then $\lambda^x, x\lambda^x, \dots, x^{m-1}\lambda^x$ are solutions of (3). Accordingly, if λ_j is a root of multiplicity m_j ($\sum_{j=1}^s m_j = n, j = 1, 2, \dots, s$), then $\lambda_j^x, x\lambda_j^x, \dots, x^{m_j-1}\lambda_j^x$ ($j = 1, \dots, s$) constitute a set of n linearly independent solutions.

Even if all the p_i are real, the characteristic equation may have complex roots. In such a case real solutions are obtained as follows: When $\lambda = \mu + iv$ is a root of multiplicity m , $\bar{\lambda} = \mu - iv$ is also a root of the same multiplicity. If we write $\rho = \sqrt{\mu^2 + v^2}$, $\tan \varphi = v/\mu$, then $\rho^x \cos \varphi x, \rho^x \sin \varphi x, x\rho^x \cos \varphi x, x\rho^x \sin \varphi x, \dots, x^{m-1}\rho^x \cos \varphi x, x^{m-1}\rho^x \sin \varphi x$ are $2m$ independent real solutions.

Nonhomogeneous equations with con-

stant coefficients can be generally solved by Lagrange's method with these solutions. However, when the nonhomogeneous term has a special form such as

$$\sum_{i=0}^n p_i y(x+i) = p(x)\lambda^x,$$

where $p(x)$ is a polynomial in x and λ is a root of multiplicity m of the characteristic equation, we can use the method of undetermined coefficients. In this particular case, the substitution $(A_0 + A_1 x + \dots + A_k x^k)x^m \lambda^x$ with undetermined coefficients A_0, A_1, \dots, A_k gives solutions, k being the degree of $p(x)$.

F. Difference and Differential Equations

The differential operator d/dx acts on the family of functions $\{x^m | m=0, \pm 1, \dots\}$ according to $dx^m/dx = mx^{m-1}$, just as the difference operator Δ acts on the family $\{x^{(m)} = \Gamma(x+1)/\Gamma(x-m+1) | m=0, \pm 1, \dots\}$ according to $\Delta x^{(m)} = mx^{(m-1)}$. Hence by using the **factorial series** $\sum a_m x^{(m)}$ and its similarity to the power series $\sum a_m x^m$, we may obtain some analogies with the theory of differential equations. For example, the \dagger Frobenius method in the theory of \dagger regular singular points can be applied to the system of difference equations

$$(z-1)\Delta_{-1} w_k(z) = \sum_{j=1}^n a_{jk}(z)w_j(z),$$

$$k=1, 2, \dots, n.$$

However, there are certain essential differences between functions defined as solutions of differential and difference equations. For example, **Hölder's theorem** states that no solution of the simple difference equation $y(x+1) - y(x) = x^{-1}$ satisfies any \dagger algebraic differential equation. Consequently, the gamma function, which is related to a solution of the equation $\psi(x) = d \log \Gamma(x)/dx$, cannot be a solution of any algebraic differential equation. For the numerical solution of ordinary differential equations by difference equation approximation — 303 Numerical Solution of Ordinary Differential Equations.

G. Geometric Difference Equations

For an arbitrary complex number q , an equation of the form $y(qx) = f(x, y(x))$ is called a **geometric difference equation**. For example, the ordinary difference equation (1) can be transformed into

$$\sum_{k=0}^n p_k(z)U(zq^k) = B(z) \quad (1')$$

by the change of variable $z = q^x$. Although it is

possible to transform an equation of the type (1') into that of the type (1), there are theories developed specifically for the type (1'), since the coefficients of the equation may become more complicated by such a transformation.

References

- [1] P. M. Batchelder, An introduction to linear difference equations, Dover, 1967.
- [2] G. Boole, A treatise on the calculus of finite differences, Dover, 1960.
- [3] R. M. Cohn, Difference algebra, Interscience, 1965.
- [4] T. Fort, Finite differences and difference equations in the real domain, Oxford Univ. Press, 1948.
- [5] C. Jordan, Calculus of finite differences, Chelsea, 1933.
- [6] K. S. Miller, Linear difference equations, Benjamin, 1968.
- [7] L. M. Milne-Thomson, The calculus of finite differences, Chelsea, 1980.
- [8] N. E. Nörlund, Vorlesungen über Differenzenrechnung, Springer, 1924.
- [9] N. E. Nörlund, Leçons sur les équations linéaires aux différences finies, Gauthier-Villars, 1929.
- [10] C. H. Richardson, An introduction to calculus of finite differences, Van Nostrand, 1954.

105 (VII.2) Differentiable Manifolds

A. General Remarks

The rudimentary concept of n -dimensional manifolds can already be seen in J. Lagrange's dynamics. In the middle of the 19th century n -dimensional \dagger Euclidean space was known as a continuum of n real parameters (A. Cayley; H. Grassmann, 1844, 1861; L. Schläfli, 1852). The notion of general n -dimensional manifolds was introduced by B. Riemann as a result of his differential geometric observations (1854). He considered an n -dimensional manifold to be a set formed by a 1-parameter family of $(n-1)$ -dimensional manifolds, just as a surface is formed by the motion of a curve. Analytical studies of topological structures of manifolds and their local properties were initiated and developed by Riemann, E. Betti, H. Poincaré, and others. To avoid the difficulties and disadvantages of analytical methods, Poincaré restricted his consideration to those topological spaces X that are \dagger connected, \dagger triangulable, and such that each point of X

has a neighborhood homeomorphic to an n -dimensional Euclidean space. We often refer to such spaces as **Poincaré manifolds**; Poincaré called them n -dimensional manifolds. In 1936, H. Whitney published a monumental paper [14] on differentiable manifolds in which the various fundamental concepts on differentiable manifolds were established. This and subsequent papers written by Whitney during nearly twenty years greatly influenced the rapid advance of the theory of differentiable manifolds since 1950.

B. Topological Manifolds

An n -dimensional **topological manifold** M is by definition a Hausdorff space in which each point p has a neighborhood $U(p)$ homeomorphic to an open set of \mathbf{R}^n .

Let M' be a Hausdorff space in which each point p of M' has a neighborhood $U(p)$ homeomorphic to an open set of H^n , where H^n is the half-space $\{(x_1, x_2, \dots, x_n) \in \mathbf{R}^n | x_n \geq 0\}$. Let $\partial M'$ denote the set consisting of points p of M' such that p corresponds to a point of $H_0^n = \{(x_1, \dots, x_n) \in H^n | x_n = 0\} \subset H^n$ under the homeomorphism from $U(p)$ to an open set of H^n . M' is called an n -dimensional **topological manifold with boundary** if $\partial M' \neq \emptyset$, and $\partial M'$ is called the **boundary** of M' . On the other hand, M defined as above or M' with $\partial M' = \emptyset$ is called an n -dimensional **topological manifold without boundary**. The **interior** of M' is the complement $M'_0 = M' - \partial M'$ of the boundary. The boundary of an n -dimensional topological manifold is an $(n-1)$ -dimensional topological manifold. A topological manifold without boundary is called **closed** or **open** according as it is compact or has no connected component which is compact. There exist connected topological manifolds that are not \dagger paracompact; among them, the 1-dimensional ones are called **long lines**. A connected paracompact topological manifold M has a \dagger countable open base and is \dagger metrizable.

C. Local Coordinates

Let M be an n -dimensional topological manifold. A pair (U, ψ) consisting of an open set U of M and a homeomorphism ψ of U onto an open set of \mathbf{R}^n is called a **coordinate neighborhood** of M . If we denote by $(x^1(p), \dots, x^n(p))$ ($p \in U$) the coordinates of the point $\psi(p)$ of \mathbf{R}^n , then x^1, x^2, \dots, x^n are real-valued continuous functions defined on U . We call these n functions the **local coordinate system** in the coordinate neighborhood (U, ψ) and the n real numbers $x^1(p), \dots, x^n(p)$ the **local coordinates** of the point $p \in U$ (with respect to (U, ψ)).

A set $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$ of coordinate neighborhoods is called an **atlas** of M if $\{U_\alpha\}_{\alpha \in A}$ forms an \dagger open covering of M .

D. Differentiable Manifolds

Let $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$ be an atlas of an n -dimensional topological manifold M . For each pair of coordinate neighborhoods (U_α, ψ_α) and (U_β, ψ_β) in S such that $U_\alpha \cap U_\beta \neq \emptyset$, $\psi_\beta \circ \psi_\alpha^{-1}$ is a homeomorphism of the open set $\psi_\alpha(U_\alpha \cap U_\beta)$ of \mathbf{R}^n onto the open set $\psi_\beta(U_\alpha \cap U_\beta)$ of \mathbf{R}^n . Let $x = (x^1, \dots, x^n) \in \psi_\alpha(U_\alpha \cap U_\beta)$. Then we can write $(\psi_\beta \circ \psi_\alpha^{-1})(x) = (f_{\beta\alpha}^1(x), \dots, f_{\beta\alpha}^n(x))$. If the n real-valued functions $f_{\beta\alpha}^1, \dots, f_{\beta\alpha}^n$ defined in $\psi_\alpha(U_\alpha \cap U_\beta)$ are of \dagger class C^r ($1 \leq r \leq \infty$) (resp. \dagger real analytic) for any α, β in A such that $U_\alpha \cap U_\beta \neq \emptyset$, then we call S an **atlas of class C^r** (resp. C^ω) of M . When an n -dimensional topological manifold M has an atlas S of class C^r ($1 \leq r \leq \omega$), we call the pair (M, S) an n -dimensional **differentiable manifold of class C^r** (or **C^r -manifold**). A C^ω -manifold is also called a **smooth manifold**, while a C^0 -manifold is called a **real analytic manifold**. We call M the **underlying topological space** of (M, S) , and we say that S defines a **differentiable structure of class C^r** (or **C^r -structure**) in M .

In particular, a C^ω -structure is called a **real analytic structure**. A C^r -manifold whose underlying topological space is compact (\dagger paracompact) is called a **compact (paracompact) C^r -manifold**. A coordinate neighborhood (U, ψ) of M is called a **coordinate neighborhood of class C^r** of (M, S) if the union $S \cup \{(U, \psi)\}$ is also an atlas of class C^r of M . In particular, each coordinate neighborhood of M belonging to S is of class C^r . The set \tilde{S} of all coordinate neighborhoods of class C^r of (M, S) is an atlas of M containing S , and we call \tilde{S} the maximal atlas containing S . Let S and S' be two atlases of class C^r of M . If $\tilde{S} = \tilde{S}'$, then we say that S and S' define the same differentiable structure of class C^r on M and that the differentiable manifolds (M, S) and (M, S') of class C^r are equivalent. In particular, (M, S) and (M, \tilde{S}) are equivalent C^r -manifolds. Let S and S' be atlases of class C^r and class C^s , respectively, where $1 \leq r < s \leq \omega$. Since $s > r$, we can consider S' an atlas of class C^r . If S and S' define the same C^r -structure in M , then we say that the C^s -structure defined by S' is **subordinate** to the C^r -structure defined by S . If M is paracompact, then there exists a C^ω -structure subordinate to a C^r -structure of M (Whitney [14]).

E. Differentiable Manifolds with Boundaries

Let U and U' be open sets in the half-space H^n , and let $\varphi: U \rightarrow U'$ be a continuous mapping. If

Differentiable Manifolds

there exist open sets W and W' in \mathbf{R}^n containing U and U' , respectively, and a mapping $\psi: W \rightarrow W'$ of class C^r that extends φ , we call φ a mapping of U into U' of class C^r . Let M be a Hausdorff topological space. A structure of a C^r -manifold on M is defined by a set $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$, where $\{U_\alpha\}_{\alpha \in A}$ is an open covering of M , and, for each α , ψ_α is a homeomorphism of U_α onto an open set of H^n such that for any $\alpha, \beta \in A$ with $U_\alpha \cap U_\beta \neq \emptyset$, $\psi_\beta \circ \psi_\alpha^{-1}$ is a mapping of class C^r from $\psi_\alpha(U_\alpha \cap U_\beta)$ onto $\psi_\beta(U_\alpha \cap U_\beta)$. Let ∂M denote the set consisting of points p of M such that $p \in U_\alpha$ and $\psi_\alpha(p) \in H_0^n = \{(x_1, \dots, x_n) \in H^n \mid x_n = 0\}$ for some $\alpha \in A$. If $\partial M \neq \emptyset$, the pair (M, S) is called an n -dimensional **differentiable manifold with boundary of class C^r** (or **C^r -manifold with boundary**), and ∂M is called the **boundary** of M . ∂M forms an $(n-1)$ -dimensional C^r -manifold. If we put $U'_\alpha = U_\alpha \cap \partial M$ and denote the restriction of ψ_α to U'_α by ψ'_α , then $S' = \{(U'_\alpha, \psi'_\alpha)\}_{\alpha \in A}$ is an atlas of class C^r of ∂M . If ∂M is empty, then (M, S) is a C^r -manifold. In this sense a C^r -manifold is sometimes called a **C^r -manifold without boundary**.

F. Orientation of a Manifold

Let $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$ be an atlas of class C^r in M , and for each α let $\{x_\alpha^1, \dots, x_\alpha^n\}$ be the local coordinate system in a coordinate neighborhood (U_α, ψ_α) . If U_α and U_β intersect, then there exist n real-valued functions F^i ($i = 1, \dots, n$) defined on $\psi_\alpha(U_\alpha \cap U_\beta)$ such that $x_\beta^i(p) = F^i(x_\alpha^1(p), \dots, x_\alpha^n(p))$ for $p \in U_\alpha \cap U_\beta$ and $i = 1, \dots, n$. The Jacobian $D_{\alpha\beta} = D(F^1, \dots, F^n)/D(x_\alpha^1, \dots, x_\alpha^n)$ is different from zero at each point $(x_\alpha^1, \dots, x_\alpha^n)$ of $\psi_\alpha(U_\alpha \cap U_\beta)$. If we can choose an atlas S of M so that, for any α, β such that $U_\alpha \cap U_\beta$ is nonempty, the Jacobian $D_{\alpha\beta}$ is always positive, then we say that the C^r -manifold M is **orientable**, and we call S an **oriented atlas**.

Let $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$ and $S' = \{(V_\lambda, \varphi_\lambda)\}_{\lambda \in \Lambda}$ be two oriented atlases of a connected C^r -manifold M . If M is connected, then the sign of the Jacobian $D_{\alpha\lambda}(p)$ of the transformation of local coordinates is independent of the choice of $\alpha \in A$, $\lambda \in \Lambda$, and $p \in U_\alpha \cap V_\lambda$. We say that S and S' define the **same (opposite) orientation** if $D_{\alpha\lambda}$ is always positive (negative). Hence if M is connected, the set of all oriented atlases of class C^r is composed of two subsets such that atlases belonging to one of them have the same orientation, while two atlases belonging to different ones have the opposite orientation. Each of these subsets is called an **orientation** of the connected C^r -manifold M . When we assign to M one of two possible orientations, M is called an **oriented manifold**; the assigned

orientation is called its **positive orientation** and the other its **negative orientation**. If $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$ belongs to the positive orientation, S and (U_α, ψ_α) are called an **atlas** and **local coordinate system**, respectively, compatible with the positive orientation.

G. Differentiable Functions

Let f be a real-valued function defined in a neighborhood of a point p of a C^∞ -manifold M . Let (U, ψ) be a coordinate neighborhood of class C^∞ such that $p \in U$. If the function $f \circ \psi^{-1}$ is of class C^r ($1 \leq r \leq \infty$) in a neighborhood of the point $\psi(p)$ in \mathbf{R}^n , then the function f is called a **function of class C^r at p** . This definition is independent of the choice of a coordinate neighborhood of class C^∞ . If we denote the local coordinate system in (U, ψ) by (x^1, \dots, x^n) , there exists a function $f(x^1, \dots, x^n)$ of n variables defined in a neighborhood of $\psi(p)$ in \mathbf{R}^n such that $f(q) = f(x^1(q), \dots, x^n(q))$ for each point q in the neighborhood of p . Here we use the same symbol f for the function f defined in a neighborhood of p in M and for the function $f \circ \psi^{-1}$ defined in the image of the neighborhood by ψ in \mathbf{R}^n . The function f is of class C^r at p if and only if $f(x^1, \dots, x^n)$ is of class C^r in a neighborhood of the point $(x^1(p), \dots, x^n(p))$ of \mathbf{R}^n . A **function of class C^r** (or **C^r -function**) **in M** is a real-valued function in M that is of class C^r at every point of M .

H. Tangent Vectors

Let M be a C^∞ -manifold, and let $\mathfrak{F}(M)$ be the real vector space consisting of all C^∞ -functions in M . (For the sake of simplicity, we denote a manifold (M, S) by M .) A **tangent vector L** at a point p of M is a linear mapping $L: \mathfrak{F}(M) \rightarrow \mathbf{R}$ such that $L(fg) = L(f)g(p) + f(p)L(g)$ for any f and g in $\mathfrak{F}(M)$. For any two tangent vectors L_1, L_2 and any pair of real numbers λ_1, λ_2 we define $\lambda_1 L_1 + \lambda_2 L_2$ by $(\lambda_1 L_1 + \lambda_2 L_2)(f) = \lambda_1 L_1(f) + \lambda_2 L_2(f)$, $f \in \mathfrak{F}(M)$.

Thus tangent vectors at p form a real vector space T_p , which we call the **tangent vector space** (or simply the **tangent space**) of M at the point p . The dimension of the tangent vector space T_p equals the dimension of M . The set of all tangent vectors of M forms a vector bundle over the base space M , called the **tangent vector bundle** (or **tangent bundle**) of M .

By a **tangent r -frame** ($r \leq n$) at p we mean an ordered set of r linearly independent tangent vectors at p . The set of all tangent r -frames also forms a fiber bundle over M called the **tangent r -frame bundle** (or **bundle of tangent r -frames**) (\rightarrow 147 Fiber Bundles F).

I. Differentials of Functions

For a C^∞ -function f in M and a point p of M we can define a linear mapping $df_p: T_p \rightarrow \mathbf{R}$ by $df_p(L) = L(f)$ for all $L \in T_p$, and we call df_p the **differential of f at p** . The totality of differentials at p of C^∞ -functions in M forms the dual vector space of the tangent vector space T_p .

J. Differentiable Mappings

Let φ be a continuous mapping of a C^∞ -manifold M into a C^∞ -manifold M' . We call φ a **differentiable mapping of class C^r** (or simply a **C^r -mapping**) ($1 \leq r \leq \infty$) if the function $f \circ \varphi$ is of class C^r for any C^r -function f on M' . If φ is a homeomorphism of M onto M' and φ and φ^{-1} are both of class C^r , then we call φ a **diffeomorphism of class C^r** . If there exists a diffeomorphism of class C^∞ of a C^∞ -manifold M onto a C^∞ -manifold M' , then M and M' are said to be **diffeomorphic**.

Let M and M' be C^∞ -manifolds and φ be a C^∞ -mapping of M into M' . For a tangent vector L of M at p , a tangent vector L' of M' at $\varphi(p)$ is defined by $L'(g) = L(g \circ \varphi)$, $g \in \mathfrak{F}(M')$. The mapping $L \rightarrow L'$ defines a linear mapping $(d\varphi)_p$ of the tangent vector space T_p of M at p into the tangent vector space $T_{\varphi(p)}$ of M' at $\varphi(p)$. The linear mapping $(d\varphi)_p$ is called the **differential of the differentiable mapping φ at p** . If $(d\varphi)_p$ is surjective, p is called a **regular point** of φ . A point on M which is not a regular point is called a **critical point** of φ . A point q on M' which is an image of a critical point is called a **critical value** of φ , and a point on M' which is not a critical value is called a **regular value**. In \mathbf{R}^n , the diffeomorphic image of a set of Lebesgue measure zero has Lebesgue measure zero. So the set of Lebesgue measure zero is well defined on a (paracompact) C^∞ -manifold. Then **Sard's theorem** states: Let $\varphi: M \rightarrow M'$ be a C^∞ -mapping; then the set of critical values of φ has Lebesgue measure zero in M' .

K. Immersions and Embeddings

Let M and M' be C^∞ -manifolds and φ be a C^∞ -mapping of M into M' . If $(d\varphi)_p$ is injective at every point p of M , then φ is called an **immersion** of M into M' . If φ is an immersion, then for some neighborhood U_p of any point p of M the restriction $\varphi|U_p$ gives rise to a homeomorphism from U_p into M' . If an immersion φ is injective, then φ is called an **embedding** (or an **imbedding**) of M into M' . An alternative definition is often used, which says that φ is an **embedding** if, in addition to the

above conditions, φ gives a homeomorphism from M onto $\varphi(M)$, where $\varphi(M)$ has the relative topology of M' . If the former definition is adopted as embedding, then a mapping φ satisfying the conditions of the alternative definition is sometimes referred to as a **regular embedding**. If M is compact, the two definitions coincide. In Sections L and M, embedding always means regular embeddings.

The theory of embeddings and immersions is mainly concerned with ways to embed and immerse a given manifold M into a manifold M' of a particular type with lowest possible dimension. M' is usually the Euclidean space \mathbf{R}^n , the projective space $P^n \mathbf{R}$ or a certain standard manifold. The theory was initiated by Whitney (1936). He proved by "general position" argument that an n -dimensional C^∞ -manifold M with countable basis can always be immersed in the $2n$ -dimensional Euclidean space and can always be embedded in the $(2n+1)$ -dimensional Euclidean space as a closed set (**Whitney's theorem**).

L. Submanifolds

A C^∞ -manifold M is said to be a **submanifold** of a C^∞ -manifold M' if M is a subset of M' and the identity mapping of M into M' is an immersion. If the identity mapping of M into M' is an embedding, then M is called a **regular submanifold** of M' . A regular submanifold M of M' is called a **closed submanifold** if M is a closed subset of M' .

Let φ be a C^∞ -mapping from M into M' and M'' be a submanifold of M' . Then for each $q \in M''$ the tangent space T_q'' of M'' at q is a linear subspace of the tangent space T_q' of M' at q . Denote by π_q the projection of quotient vector space onto T_q'/T_q'' . A C^∞ -mapping φ is called **transverse** to M'' if for each $p \in \varphi^{-1}(M'')$ the composite $\pi_{\varphi(p)} \circ d\varphi_p: T_p \rightarrow T_{\varphi(p)}'/T_{\varphi(p)}''$ is surjective. If φ is transverse to M'' then $\varphi^{-1}(M'')$ is a submanifold of M . For any C^∞ -mapping $\varphi: M \rightarrow M'$ and any submanifold $M'' \subset M'$, we can find a C^∞ -mapping $\varphi': M \rightarrow M'$ which is transverse to M'' and arbitrarily close to φ (**transversality theorem**). Let M_1 and M_2 be submanifolds of M' . Then we say M_1 **intersects transversely** to M_2 if the inclusion $M_1 \subset M'$ is transverse to M_2 .

A C^∞ -mapping is called a **submersion** if it has no critical point. Let φ be a submersion from M into M' . Then for each point $q \in M'$, $\varphi^{-1}(q)$ is a regular submanifold of M , and M is covered by a mutually disjoint family of submanifolds: $M = \bigcup_{q \in M'} \varphi^{-1}(q)$.

Let M be a submanifold of an n -dimensional Euclidean space \mathbf{R}^n . We can identify the tangent vector space T_p of M at p with the geom-

etric tangent space of M at p in the Euclidean space \mathbf{R}^m . A vector in \mathbf{R}^m that is orthogonal to the tangent vector space T_p of M at p is called a **normal vector** to M at p . The set of all vectors normal to M forms a vector bundle over M , which we call the **normal vector bundle** (or **normal bundle**) of M . If M is compact, then the totality of vectors normal to M whose length is $\leq \varepsilon$ (where ε is a sufficiently small positive real number) forms a neighborhood $N(M)$ of M in \mathbf{R}^m which we call a **tubular neighborhood** of M . $\text{Int } N(M)$ is called an **open tubular neighborhood**.

M. Vector Fields

Let N be a subset of a C^∞ -manifold M . By a **vector field** on N we mean a mapping X that assigns to each point p of N a tangent vector X_p of M at p . We can consider X as a cross section over N of the tangent vector bundle of M . Let X be a vector field in M , and let f be a C^∞ -function in M . Then we can define a function Xf in M by $(Xf)(p) = X_p f$. We call X a **vector field of class C^r** if the function Xf is of class C^r for any C^∞ -function f in M . Let (x^1, \dots, x^n) be the local coordinate system in a coordinate neighborhood (U, ψ) , and let $(\partial/\partial x^i)_p f = (\partial f/\partial x^i)(p)$ for $p \in U$ and $f \in \mathfrak{F}(M)$. Then $\partial/\partial x^i$ ($i = 1, \dots, n$) are vector fields in U , and the $(\partial/\partial x^i)_p$ form a basis of T_p at every point $p \in U$. A vector field X in U is written uniquely as $X_p = \sum_i \xi^i(p) (\partial/\partial x^i)_p$ at each point $p \in U$. Then ξ^1, \dots, ξ^n are real-valued functions defined in U , called the **components** of X with respect to the local coordinate system (x^1, \dots, x^n) . A vector field X in M is of class C^r if and only if its components ξ^i with respect to each coordinate system are functions of class C^r ($0 \leq r \leq \infty$). Let $(\bar{x}^1, \dots, \bar{x}^n)$ be another local coordinate system in a neighborhood U of p , and let $(\bar{\xi}^1, \dots, \bar{\xi}^n)$ be the components of X with respect to $(\bar{x}^1, \dots, \bar{x}^n)$. Then we have $\bar{\xi}^i(q) = \sum_j (\partial \bar{x}^i/\partial x^j)(q) \xi^j(q)$ at each point $q \in U$.

For the rest of this article we mean by a vector field in M a vector field of class C^∞ , and we denote by $\mathfrak{X}(M)$ the set of all vector fields in M . Then $\mathfrak{X}(M)$ is an $\mathfrak{F}(M)$ - $\mathfrak{X}(M)$ -module, where $\mathfrak{F}(M)$ denotes the algebra of all C^∞ -functions in M . In fact, for $f, g \in \mathfrak{F}(M)$ and $X, Y \in \mathfrak{X}(M)$, we can define a vector field $fX + gY$ by $(fX + gY)_p = f(p)X_p + g(p)Y_p$, and this defines an $\mathfrak{F}(M)$ -module structure in $\mathfrak{X}(M)$.

In a coordinate neighborhood (U, ψ) , we can write $X = \sum_i \xi^i (\partial/\partial x^i)$. The right-hand side of this equation is sometimes called the **symbol** of the vector field X . A vector field X can also be interpreted as a linear differential operator that acts on $\mathfrak{F}(M)$.

Let X and Y be vector fields in M . Then

there exists a unique vector field Z in M such that $Zf = X(Yf) - Y(Xf)$ for any C^∞ -function f in M . We denote Z by $[X, Y]$ and call it the **Poisson bracket** (or simply **bracket**) of X and Y . If ξ^i and η^i denote the components of X and Y , respectively, in a coordinate neighborhood (U, ψ) , then the components ζ^i of $[X, Y]$ are given by $\zeta^i = \sum_k \{ \xi^k (\partial \eta^i/\partial x^k) - \eta^k (\partial \xi^i/\partial x^k) \}$. The bracket of vector fields has the following properties: (i) $[X, Y]f = X(Yf) - Y(Xf)$, (ii) $[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X$, (iii) $[X + Y, Z] = [X, Z] + [Y, Z]$, (iv) $[X, Y] = -[Y, X]$, and (v) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (**Jacobi identity**). These identities show that $\mathfrak{X}(M)$ is a Lie algebra (\rightarrow 248 Lie Algebras) over \mathbf{R} .

If φ is a diffeomorphism of M onto M' , then for any vector field X in M we can define a vector field $\varphi_* X$ in M' by the condition $(\varphi_* X)_p = d\varphi_q(X_q)$, $p = \varphi(q)$. Then φ_* is an isomorphism of the Lie algebra $\mathfrak{X}(M)$ onto the Lie algebra $\mathfrak{X}(M')$.

N. Vector Fields and One-Parameter Groups of Transformations

A **one-parameter group of transformations** of M is a family φ_t ($t \in \mathbf{R}$) of diffeomorphisms satisfying the following two conditions: (i) the mapping of $\mathbf{R} \times M$ into M defined by $(t, p) \rightarrow \varphi_t(p)$ is of class C^∞ ; and (ii) $\varphi_s \circ \varphi_t = \varphi_{s+t}$ for $s, t \in \mathbf{R}$.

Let φ_t be a one-parameter group of transformations of M . Then we can define a vector field X by $X_p f = \lim_{t \rightarrow 0} (f(\varphi_t(p)) - f(p))/t$, where $p \in M$ and $f \in \mathfrak{F}(M)$. The vector field X thus defined is called the **infinitesimal transformation** of φ_t . We also say that φ_t is generated by X , and sometimes we denote φ_t by the symbol $\exp tX$. In this case, if (x^1, \dots, x^n) is a local coordinate system, then at each point p of the coordinate neighborhood, we have $X_p = \sum_i (dx^i(\varphi_t(p))/dt)_{t=0} (\partial/\partial x^i)_p$.

If M is compact, then every vector field in M is the infinitesimal transformation of a one-parameter group of transformations; that is, every vector field generates a one-parameter group of transformations. For M noncompact, this is not always true. Nevertheless, for each vector field X we have the following result concerning local properties of X : For each point p of M , there exist a neighborhood U of p , a positive real number ε , and a family φ_t ($|t| < \varepsilon$) of mappings of U into M satisfying the following three conditions. (1) The mapping of $(-\varepsilon, \varepsilon) \times U$ into M defined by $(t, q) \rightarrow \varphi_t(q)$ is of class C^∞ , and for each fixed t , φ_t is a diffeomorphism of U onto an open set $\varphi_t(U)$ of M . (2) If $|s|$, $|t|$, and $|s+t|$ are all smaller than ε and q and $\varphi_t(q)$ both belong to U , then

$\varphi_s(\varphi_t(q)) = \varphi_{s+t}(q)$. (3) $X(q)f = \lim_{t \rightarrow 0} (f(\varphi_t(q)) - f(q))/t$ for $q \in U$ and $f \in \mathfrak{F}(M)$. We call φ_t the **local one-parameter group of local transformations** around p generated by X .

Let X and Y be vector fields in M , and let φ_t be the local one-parameter group of local transformations around p generated by X . Then $[X, Y]_p = \lim_{t \rightarrow 0} (Y_p - ((\varphi_t)_* Y)_p)/t$, where $(\varphi_t)_* Y$ is a vector field defined as follows: Let U be a neighborhood of the point p where φ_t ($|t| < \varepsilon$) is defined. Then $(\varphi_t)_* Y$ is the vector field on $\varphi_t(U)$ that is the image of Y under the diffeomorphism φ_t . In particular, if X generates a one-parameter group of transformations of M , then we have $[X, Y] = \lim_{t \rightarrow 0} (Y - (\varphi_t)_* Y)/t$ for any vector field Y in M .

O. Tensor Fields

Let $T_s^r(p)$ be the vector space consisting of all r -times contravariant and s -times covariant tensors over the tangent vector space T_p of a C^∞ -manifold M , that is,

$$T_s^r(p) = \left(\bigotimes^r T_p \right) \otimes \left(\bigotimes^s T_p^* \right),$$

where T_p^* denotes the dual linear space of T_p (\rightarrow 256 Linear Spaces). A **tensor field** (more precisely, **contravariant of order r and covariant of order s** , or simply a **tensor field of type (r, s)**) on a subset N of M is a mapping K that assigns to each point p of N an element K_p of the vector space $T_s^r(p)$. In particular, if $r = s = 0$, K is a real-valued function on N , and we call K a **scalar field**. If $r = 1$ and $s = 0$, K is a vector field, called a **contravariant vector field**. When $r = 0$ and $s = 1$, we call K a **covariant vector field** (or **differential form of degree 1**). If $r \neq 0$, $s = 0$ or $r = 0$, $s \neq 0$, we call K a **contravariant tensor field of order r** or a **covariant tensor field of order s** , respectively. A contravariant or covariant tensor field K is said to be **symmetric (alternating)** if K_p is a symmetric (alternating) tensor at every point p of M .

Let (x^1, \dots, x^n) be the local coordinate system in a coordinate neighborhood (U, ψ) . Then at each point p of U , the $(\partial/\partial x^i)_p$ ($i = 1, \dots, n$) form a basis of the tangent vector space T_p , the differentials $(dx^i)_p$ ($i = 1, \dots, n$) form a basis of the dual space T_p^* , and these bases are dual to each other. A tensor field K of type (r, s) defined on M is written at any point p of U in the following form:

$$K_p = \sum K_{j_1 \dots j_s}^{i_1 \dots i_r}(p) (\partial/\partial x^{i_1})_p \otimes \dots \otimes (\partial/\partial x^{i_r})_p \\ \otimes (dx^{j_1})_p \otimes \dots \otimes (dx^{j_s})_p.$$

The functions $K_{j_1 \dots j_s}^{i_1 \dots i_r}$ defined in U are called the **components** of the tensor field K of type (r, s) with respect to the local coordinate sys-

Differentiable Manifolds

tem (x^1, \dots, x^n) . If $\bar{K}_{j_1 \dots j_s}^{i_1 \dots i_r}$ are the components of K with respect to the local coordinate system $(\bar{x}^1, \dots, \bar{x}^n)$ in another coordinate neighborhood (U', ψ') such that $U \cap U' \neq \emptyset$, then for each $q \in U \cap U'$, the following relations hold:

$$\bar{K}_{j_1 \dots j_s}^{i_1 \dots i_r}(q) = \sum_{k_1, \dots, k_r} (\partial \bar{x}^{i_1} / \partial x^{k_1})_q \dots (\partial \bar{x}^{i_r} / \partial x^{k_r})_q \\ \times (\partial x^{j_1} / \partial \bar{x}^{j_1})_q \dots (\partial x^{j_s} / \partial \bar{x}^{j_s})_q K_{j_1 \dots j_s}^{k_1 \dots k_r}.$$

A tensor field K in M is called a **tensor field of class C^t** ($0 \leq t \leq \infty$) if the components are functions of class C^t for any coordinate neighborhood of M .

The sum $K + L$ and the tensor product $K \otimes L$ of two tensor fields K and L are defined by the rules $(K + L)_p = K_p + L_p$ and $(K \otimes L)_p = K_p \otimes L_p$, respectively. The contraction of two tensor fields is also defined by taking the contraction pointwise.

Let φ be a diffeomorphism of M into M' . Then the differential $(d\varphi)_q$ is an isomorphism of T_q onto T_p ($p = \varphi(q)$) for each $q \in M$ and hence induces an isomorphism $\tilde{\varphi}_q$ of the vector space $T_s^r(q)$ onto the vector space $T_s^r(p)$ (\rightarrow 256 Linear Spaces). For any tensor field K in M we can define a tensor field $\tilde{\varphi}K$ in M' by $(\tilde{\varphi}K)_p = \tilde{\varphi}_q(K_q)$, $p = \varphi(q)$, $q \in M$. Then $\tilde{\varphi}(K + L) = \tilde{\varphi}K + \tilde{\varphi}L$, $\tilde{\varphi}(K \otimes L) = \tilde{\varphi}K \otimes \tilde{\varphi}L$, and the mapping $\tilde{\varphi}$ commutes with contraction.

Let K be a tensor field and X be a vector field (both of class C^∞) in M . We define a tensor field $L_X K$ by $(L_X K)_p = \lim_{t \rightarrow 0} (K_p - (\tilde{\varphi}_t K)_p)/t$, where φ_t denotes the local one-parameter group of local transformations around p generated by X . We call $L_X K$ the **Lie derivative** of K with respect to the vector field X . The operator $L_X: K \rightarrow L_X K$ has the following six properties: (i) $L_X(K + K') = L_X K + L_X K'$; (ii) $L_X(K \otimes K') = (L_X K) \otimes K' + K \otimes L_X K'$; (iii) the operator L_X commutes with contraction; (iv) $L_X f = Xf$ for a scalar field f and $L_X Y = [X, Y]$ for a vector field Y ; (v) $L_{[X, Y]} = L_X L_Y - L_Y L_X$, that is, $L_{[X, Y]} K = L_X(L_Y K) - L_Y(L_X K)$; and (vi) K is invariant under φ_t , i.e., $\tilde{\varphi}_t K = K$ for all t , if and only if $L_X K = 0$.

Let K be a covariant tensor field of order r in M . We always assume that K is of class C^∞ . The value K_p of K at $p \in M$ is an element of the vector space $T_p^* \otimes \dots \otimes T_p^*$ (r times tensor product of T_p^*); hence we may consider K_p an r -linear mapping of T_p into \mathbf{R} (\rightarrow 256 Linear Spaces). If X_1, \dots, X_r are vector fields in M , we define a C^∞ -function $K(X_1, \dots, X_r)$ by $K(X_1, \dots, X_r)(p) = K_p((X_1)_p, \dots, (X_r)_p)$. Then the mapping that assigns to each r -tuple (X_1, \dots, X_r) of vector fields the C^∞ -function $K(X_1, \dots, X_r)$ is an r -linear mapping on the $\mathfrak{F}(M)$ -module $\mathfrak{X}(M)$ consisting of all vector fields of class C^∞ in M into $\mathfrak{F}(M)$; that is, $K(X_1, \dots, fX_i + gY_i, \dots, X_r) = fK(X_1, \dots, X_i,$

$\dots, X_r) + gK(X_1, \dots, Y_i, \dots, X_r)$ ($i = 1, \dots, r$) for $f, g \in \mathfrak{F}(M)$. Conversely any r -linear mapping of the $\mathfrak{F}(M)$ -module $\mathfrak{X}(M)$ into $\mathfrak{F}(M)$ can be interpreted as a covariant tensor field of order r in M . If the tensor field K is symmetric (alternating), the corresponding r -linear mapping $K(X_1, \dots, X_r)$ is symmetric (alternating) with respect to X_1, \dots, X_r . For the Lie derivative $L_X K$ of a covariant tensor field of order r of K , we have the following formula: $(L_X K)(X_1, \dots, X_r) = X(K(X_1, \dots, X_r)) - \sum_{i=1}^r K(X_1, \dots, [X, X_i], \dots, X_r)$.

P. Riemannian Metrics

A symmetric covariant tensor field g of order 2 and of class C^∞ in M is called a **pseudo-Riemannian metric** if the symmetric bilinear form g_p on the tangent vector space T_p is non-degenerate at each point $p \in M$; and g is called a **Riemannian metric** if g_p is positive definite for all p . If g is a Riemannian metric, the length $\|L\|$ of a tangent vector $L \in T_p$ is defined by $\|L\|^2 = g_p(L, L)$. On a paracompact C^∞ -manifold there always exists a Riemannian metric. A pair consisting of a differentiable manifold and a Riemannian metric on it is called a **Riemannian manifold** (\rightarrow 364 Riemannian Manifolds).

Q. Differential Forms

An alternating covariant tensor field in M of order r and of class C^t ($0 \leq t \leq \infty$) is also called a **differential form** (or **exterior differential form**) of degree r . A differential form of degree 1 is sometimes called a **Pfaffian form**. Let ω be a differential form of degree r . Since each alternating covariant tensor of order r at a point p is an element of $\wedge^r T_p^*$, the r -fold exterior product of T_p^* , the form ω is a mapping that sends each point p of M to an element ω_p of $\wedge^r T_p^*$. We can also regard ω as an alternating r -linear mapping of $\mathfrak{X}(M)$ into $\mathfrak{F}(M)$. Let (x^1, \dots, x^n) be the local coordinate system in a local coordinate neighborhood (U, ψ) . Since $(dx^i)_p$ ($i = 1, \dots, n$) is a basis of T_p^* at each point p of U , we can express ω_p ($p \in U$) uniquely in the form

$$\omega_p = \sum_{i_1 < \dots < i_r} a_{i_1 \dots i_r}(p) (dx^{i_1})_p \wedge \dots \wedge (dx^{i_r})_p,$$

where the sum extends over all ordered r -tuples (i_1, \dots, i_r) of indices such that $1 \leq i_1 < i_2 < \dots < i_r \leq n$. For an ordered r -tuple (i_1, \dots, i_r) of indices with repeated indices we put $a_{i_1 \dots i_r} = 0$, and for (i_1, \dots, i_r) with r distinct indices we put $a_{i_1 \dots i_r} = (\text{sgn } \sigma) a_{j_1 \dots j_r}$, where (j_1, \dots, j_r) ($j_1 < \dots < j_r$) is a permutation of (i_1, \dots, i_r) and $\text{sgn } \sigma$ denotes the sign of the permutation $\sigma: i_k \rightarrow j_k$

($k = 1, \dots, r$). Then we can write

$$\omega_p = \frac{1}{r!} \sum_{i_1, \dots, i_r=1}^n a_{i_1 \dots i_r}(p) (dx^{i_1})_p \wedge \dots \wedge (dx^{i_r})_p.$$

The functions $a_{i_1 \dots i_r}$ are the components of the tensor field ω , and ω is of class C^t if these components are of class C^t for any coordinate neighborhood. By the **support** (or **carrier**) of a differential form ω we mean the closure of the subset of M consisting of all p such that $\omega_p \neq 0$.

In the rest of this article, differential forms are always of class C^∞ , and we denote by $\mathfrak{D}^r(M)$ the real vector space consisting of all differential forms of degree r and of class C^∞ . In particular, $\mathfrak{D}^0(M) = \mathfrak{F}(M)$ and $\mathfrak{D}^r(M) = \{0\}$ for $r > n$, $n = \dim M$.

For differential forms we have the following five important operations.

(1) Exterior product. Let ω and θ be differential forms of degree r and s , respectively. The **exterior product** $\omega \wedge \theta$ of ω and θ is the differential form of degree $r+s$ defined by $(\omega \wedge \theta)_p = \omega_p \wedge \theta_p$, $p \in M$. Let X_1, \dots, X_{r+s} be $r+s$ vector fields in M . Then we have

$$(\omega \wedge \theta)(X_1, \dots, X_{r+s}) = \sum \text{sgn}(i; j) \omega(X_{i_1}, \dots, X_{i_r}) \theta(X_{j_1}, \dots, X_{j_s}),$$

where the summation runs over all possible partitions of $(1, 2, \dots, r+s)$ such that $i_1 < i_2 < \dots < i_r$, and $j_1 < j_2 < \dots < j_s$, and $\text{sgn}(i; j)$ means the sign of the permutation $(1, 2, \dots, r+s) \rightarrow (i_1, \dots, i_r, j_1, \dots, j_s)$. In particular, if $\omega_1, \dots, \omega_r$ are differential forms of degree 1, then we have $(\omega_1 \wedge \dots \wedge \omega_r)(X_1, \dots, X_r) = \det(\omega_i(X_j))$.

(2) Exterior differentiation. Let ω be a differential form of degree r , and let $\omega = (1/r!) \sum a_{i_1 \dots i_r} dx^{i_1} \wedge \dots \wedge dx^{i_r}$ in a coordinate neighborhood (U, ψ) . Then we can define a differential form $d\omega$ of degree $r+1$ by the condition $d\omega = (1/r!) \sum da_{i_1 \dots i_r} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_r}$ in U . The differential form $d\omega$ is called the **exterior derivative** (or **exterior differential**) of ω . A differential form ω satisfying the condition $d\omega = 0$ is called a **closed differential form**, and a differential form η that can be expressed as $\eta = d\omega$ for some ω is called an **exact differential form**. If $a, b \in \mathbf{R}$ and $\omega, \omega' \in \mathfrak{D}^r(M)$, then we have $d(a\omega + b\omega') = ad\omega + bd\omega'$. Therefore the set $\mathfrak{C}^r(M)$ of all closed differential forms of degree r and the set $\mathfrak{E}^r(M)$ of all exact differential forms of degree r are linear subspaces of $\mathfrak{D}^r(M)$.

For the exterior derivative $d\omega$, we have the following formula:

$$\begin{aligned} (d\omega)(X_1, \dots, X_r) &= \sum_{i=1}^{r+1} (-1)^{i+1} X_i(\omega(X_1, \dots, \hat{X}_i, \dots, X_{r+1})) \\ &\quad + \sum_{i < j} (-1)^{i+j} \omega([X_i X_j], X_1, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_{r+1}), \end{aligned}$$

where the variables under the sign \wedge are to be omitted.

(3) Interior product with vector field. Let ω be a differential form of degree r and X a vector field. When $r \geq 1$, we can define a differential form $\iota(X)\omega$ of degree $r-1$ by the formula $(\iota(X)\omega)(X_1, \dots, X_{r-1}) = \omega(X, X_1, \dots, X_{r-1})$ for any $r-1$ vector fields X_1, \dots, X_{r-1} ; if $r=0$, we put $\iota(X)\omega=0$. The differential form $\iota(X)\omega$ is called the **interior product** of ω with X .

(4) Lie derivative. The **Lie derivative** $L_X\omega$ of a differential form of degree r with respect to a vector field X is a differential form of the same degree. For any r vector fields X_1, \dots, X_r , we have, by definition, $(L_X\omega)(X_1, \dots, X_r) = X(\omega(X_1, \dots, X_r)) - \sum_{i=1}^r \omega(X_1, \dots, [X, X_i], \dots, X_r)$.

(5) Let φ be a C^∞ -mapping of M into M' , and let $\varphi_p^*: T_{\varphi(p)}^* \rightarrow T_p^*$ be the \dagger transpose (or \dagger dual) of the linear mapping $(d\varphi)_p: T_p \rightarrow T_{\varphi(p)}$ for $p \in M$, i.e., the mapping defined by the condition $((d\varphi)_p L, \alpha) = (L, \varphi_p^* \alpha)$ for each $L \in T_p$, $\alpha \in T_{\varphi(p)}^*$. We denote the linear mapping of $\wedge T_{\varphi(p)}^*$ into $\wedge T_p^*$ induced by φ_p^* by φ_p^* also. Let ω be a differential form of degree r in M . Then a differential form $\varphi^*\omega$ in M , the **pullback** by φ of ω , is defined by $(\varphi^*\omega)_p = \varphi_p^* \omega_{\varphi(p)}$, $p \in M$.

The operations defined previously satisfy the following six important relations: (i) $d^2 = 0$, that is, $d(d\omega) = 0$; (ii) $d(\omega \wedge \theta) = d\omega \wedge \theta + (-1)^r \omega \wedge d\theta$, where ω is of degree r ; (iii) $\varphi^*(\omega \wedge \theta) = \varphi^*\omega \wedge \varphi^*\theta$, $\varphi^*(d\omega) = d(\varphi^*\omega)$; (iv) $L_X(\omega \wedge \theta) = (L_X\omega) \wedge \theta + \omega \wedge (L_X\theta)$; (v) $L_X = \iota(X) \cdot d + d \cdot \iota(X)$, $L_X(d\omega) = d(L_X\omega)$; and (vi) $L_{[X, Y]} = L_X \cdot L_Y - L_Y \cdot L_X$, $\iota([X, Y]) = L_X \cdot \iota(Y) - \iota(Y) \cdot L_X$.

R. De Rham Cohomology

Let $\mathfrak{D}(M) = \sum_{r=0}^n \mathfrak{D}^r(M)$, where $n = \dim M$. Then $\mathfrak{D}(M)$ is a \dagger cochain complex with \dagger coboundary operator d . We denote by $H^r(\mathfrak{D})$ the r -dimensional cohomology group of this cochain complex, and we call it the r -dimensional **de Rham cohomology group** of the differentiable manifold M . If we denote by $\mathfrak{C}^r(M)$ and $\mathfrak{E}^r(M)$ the subspaces of $\mathfrak{D}^r(M)$ consisting of closed differential forms and exact differential forms, respectively, then $H^r(\mathfrak{D}) = \mathfrak{C}^r(M)/\mathfrak{E}^r(M)$ ($0 \leq r \leq n$) by definition. If $\omega \in \mathfrak{C}^i(M)$ and $\theta \in \mathfrak{E}^r(M)$, then $\omega \wedge \theta \in \mathfrak{E}^{i+r}(M)$, and if $\omega \in \mathfrak{C}^i(M)$ and $\theta \in \mathfrak{E}^r(M)$ (or $\omega \in \mathfrak{E}^i(M)$ and $\theta \in \mathfrak{E}^r(M)$), then $\omega \wedge \theta \in \mathfrak{E}^{i+r}(M)$. So if we put $H(\mathfrak{D}) = \sum_{r=1}^n H^r(\mathfrak{D})$ (direct sum), we can define a product in $H(\mathfrak{D})$ by $[\omega] \cdot [\theta] = [\omega \wedge \theta]$ for each $[\omega] \in H^i(\mathfrak{D})$, $[\theta] \in H^r(\mathfrak{D})$. With respect to this product, $H(\mathfrak{D})$ forms an algebra over \mathbf{R} called the **de Rham cohomology ring** of M .

S. Partitions of Unity

Let M be a paracompact C^∞ -manifold, and let $\{V_i\}_{i \in I}$ be a \dagger locally finite open covering of M such that the closure of V_i is compact for each index i . Then there exists a C^∞ -function f_i in M for each i satisfying the following three conditions: (i) $0 \leq f_i \leq 1$, (ii) the support of f_i is contained in V_i , and (iii) $\sum_{i \in I} f_i(x) = 1$ for every $x \in M$. The family of C^∞ -functions f_i , $i \in I$, is called a **partition of unity of class C^∞** subordinate to the open covering $\{V_i\}_{i \in I}$.

T. Integrals of Differential Forms

Integrals over an Oriented Manifold. Let M be an n -dimensional paracompact oriented C^∞ -manifold and ω a differential form of degree n in M with compact support. We can choose a positively oriented atlas $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$ such that $\{U_\alpha\}_{\alpha \in A}$ is a locally finite covering; \bar{U}_α is compact for each α . Suppose first that the support of ω is contained in U_α for some index α . Then $(\psi_\alpha^{-1})^* \omega_\alpha$ is a differential form of degree n in $\psi_\alpha(U_\alpha)$, and we can express $(\psi_\alpha^{-1})^* \omega_\alpha$ in the form $adx^1 \wedge \dots \wedge dx^n$, where (x^1, \dots, x^n) are the coordinates in \mathbf{R}^n such that $x_\alpha^i = x^i \circ \psi_\alpha$ ($i = 1, \dots, n$) give a local coordinate system compatible with the orientation of M and a is a C^∞ -function with compact support. Then we define the integral of ω over M by

$$\int_M \omega = \int_{\psi_\alpha(U_\alpha)} a dx^1 \dots dx^n.$$

For the general case let $\{f_\alpha\}_{\alpha \in A}$ be a partition of unity of class C^∞ subordinate to $\{U_\alpha\}_{\alpha \in A}$. Then the support of $f_\alpha \omega$ is contained in U_α , and except for a finite number of the indices α , $f_\alpha \omega$ vanishes identically. Therefore we may define the integral of ω over M by

$$\int_M \omega = \sum_\alpha \int_M f_\alpha \omega,$$

and we can show that this definition of the integral is independent of the choice of oriented atlas S and of a partition of unity subordinate to S .

Integrals over a Singular Chain. We fix rectangular coordinates in \mathbf{R}^r . Let d_0 be the origin and d_i be the unit point on the i th coordinate axis. Let S^r denote the oriented r -simplex (d_0, d_1, \dots, d_r) with vertices d_0, d_1, \dots, d_r . When we regard S^r as a point set, we denote it by $|S^r|$. An **oriented singular r -simplex of class C^∞** in M is, by definition, a pair (S^r, φ) consisting of S^r and a C^∞ -mapping φ of an open neighborhood of $|S^r|$ into M . An element of the free \mathbf{Z} -module generated by singular r -simplexes of class C^∞ is called an **integral singular r -chain**

Differentiable Manifolds

of class C^∞ in M . We define a real singular r -chain of class C^∞ analogously. Let ω be a differential form of degree r and (S^r, φ) be an oriented singular r -simplex of class C^∞ in M . Then $\varphi^*\omega$ is a differential form of degree r in a neighborhood of $|S^r|$, and we can express $\varphi^*\omega$ in the form $\varphi^*\omega = a dx^1 \wedge dx^2 \wedge \dots \wedge dx^r$. We define the integral of ω over (S^r, φ) by

$$\int_{(S^r, \varphi)} \omega = \int_{|S^r|} a dx^1 \dots dx^r,$$

and the integral of ω over a singular r -chain C of class C^∞ by

$$\int_C \omega = \sum_i m_i \int_{(S^r, \varphi_i)} \omega,$$

where $C = \sum_i m_i (S^r, \varphi_i)$, $m_i \in \mathbf{Z}$ (or $m_i \in \mathbf{R}$).

When $r=0$, then ω is a function in M , and S^0 is a point o . In this case we put $\int_C \omega = \sum_i m_i \omega(\varphi_i(o))$.

Let $C_r(S, \mathbf{Z})$ ($C_r(S, \mathbf{R})$) be the \mathbf{Z} -module (vector space over \mathbf{R}) of integral (real) singular r -chains of class C^∞ in M , and let $\omega(C)$ be the value of the integral of ω over a chain C . Then ω is a linear function in the vector space $C_r(S, \mathbf{R})$, and hence we can consider ω a **singular r -cochain of class C^∞** .

U. Stokes's Formulas

(1) Let D be a domain in an n -dimensional C^∞ -manifold M , and let ∂D and \bar{D} be the boundary and the closure of D , respectively. Let $S = \{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$ be an atlas of class C^∞ of M , $U'_\alpha = U_\alpha \cap \bar{D}$, ψ'_α be the restriction of ψ_α to U'_α , and $T = \{(U'_\alpha, \psi'_\alpha)\}_{\alpha \in A}$. If the pair (\bar{D}, T) is a C^∞ -manifold with boundary under a suitable choice of S , then the domain D is called a **domain with regular (or smooth) boundary**.

The boundary ∂D of (\bar{D}, T) is then an $(n-1)$ -dimensional closed submanifold of M , and if M is orientable, ∂D is also orientable. Now let M be a paracompact and oriented manifold and D be a domain with regular boundary. Let C be a characteristic function of D in M , i.e., a function defined by the condition $C(p) = 1$ for $p \in D$ and $C(p) = 0$ for $p \notin D$. Let θ be a differential form of degree n in M with compact support. We define the integral of θ over D by

$$\int_D \theta = \int_M C \cdot \theta.$$

Let ω be a differential form of degree $n-1$ in M with compact support. We then have **Stokes's formula**:

$$\int_D d\omega = \int_{\partial D} i^* \omega,$$

where i denotes the identity mapping of the

submanifold ∂D into M with ∂D having the orientation induced naturally from that of M .

(2) Let C be a singular r -chain of class C^∞ in M , and let ∂C be the boundary of C . Then for any differential form ω of degree $r-1$, we have

$$\int_C d\omega = \int_{\partial C} \omega.$$

This formula is also called **Stokes's formula**.

V. De Rham's Theorem

Let M be a connected paracompact C^∞ -manifold. If we consider ω as a singular cochain, we have $(d\omega)(C) = \omega(\partial C)$; by Stokes's formula, this means that the exterior differential $d\omega$ of ω is equal to the coboundary of the singular cochain ω . Let ω and C be a closed differential form of degree r and a singular r -cycle of class C^∞ , respectively, and let $[\omega]$ and $[C]$ be the de Rham cohomology class and the singular homology class represented by ω and C . Using Stokes's formula, we can define the inner product $([\omega], [C])$ by

$$([\omega], [C]) = \int_C \omega.$$

Through this inner product, it follows that the de Rham cohomology group $H^r(\mathfrak{D})$ is isomorphic to the r th singular cohomology group $H^r(M, \mathbf{R})$, the dual space of the r th homology group of the complex of real singular chains of class C^∞ . Moreover, the de Rham cohomology ring $H(\mathfrak{D})$ is isomorphic to the singular cohomology ring $H^*(M, \mathbf{R})$ (**de Rham's theorem**).

W. Divergence of a Vector Field

Let M be an n -dimensional oriented C^∞ -manifold, and let S be an oriented atlas. Let ω be a differential form of degree n , and let (x^1, \dots, x^n) be the local coordinate system in a coordinate neighborhood in S . Then we can express ω in the coordinate neighborhood uniquely in the form $\omega = a dx^1 \wedge \dots \wedge dx^n$. If the function a is positive for any coordinate neighborhood in S , we call ω a **volume element** of M . In a paracompact oriented manifold, there always exists a volume element. (We remark that an n -dimensional differentiable manifold M is orientable if and only if there exists an everywhere nonvanishing differential form of degree n .) Let f be a C^∞ -function in M with compact support. Then $f \cdot \omega$ is a differential form of degree n with compact support, and so the integral

$$\int_M f \cdot \omega$$

is defined. We call this integral the **integral of the function f with respect to the volume element ω** .

Let g be a Riemannian metric in M and g_{ij} the components of g with respect to the local coordinate system (x^1, \dots, x^n) as before. Then we can define a volume element ω in M by putting $\omega = \sqrt{G} dx^1 \wedge \dots \wedge dx^n$, $G = \det(g_{ij})$ in each coordinate neighborhood. The volume element thus defined is called the **volume element associated with the Riemannian metric g** .

Let ω be a volume element and X a vector field in M . Then the Lie derivative $L_X \omega$ is also a differential form of degree n , and we can express $L_X \omega$ in the form $L_X \omega = f_X \cdot \omega$, where f_X is a scalar field, i.e., a function in M . We call f_X the **divergence** of the vector field X with respect to the volume element ω and denote it by $\text{div } X$. If ω is associated with a Riemannian metric, then $\text{div } X$ is called the **divergence** of X with respect to the Riemannian metric.

If M is compact, we have

$$\int_M \text{div } X \cdot \omega = 0$$

for any vector field X . This result is known as **Green's theorem**.

X. Jets

Let M and N be C^∞ -manifolds. We define an equivalence relation in the set of all C^∞ -mappings of M into N . Let f and g be such mappings and p be a point of M . Choosing local coordinate systems, we write $f(p) = (f_1(x), \dots, f_n(x))$, $g(p) = (g_1(x), \dots, g_n(x))$, $x = (x_1, \dots, x_m)$. We say f and g are equivalent at p if $f(p)$, $g(p)$, and the values at p of all the partial derivatives of f_i and g_i up to the order r (r an integer, $r \geq 0$) are equal ($i = 1, \dots, n$). An equivalence class with respect to this equivalence relation is called a **jet of order r at p** . A jet of order r at p represented by a function f is denoted by $j_p^r f$, and the points p and $f(p)$ are called the **source** and the **target** of the jet $j_p^r f$, respectively. We denote by $J_p^r(M, N)$ the set of all jets of order r with source at p and target in N and let $J^r(M, N) = \bigcup_{p \in M} J_p^r(M, N)$. For any jet j , let $\pi_s(j)$ and $\pi_t(j)$ denote the source and the target of j , respectively. We can introduce the structure of a C^∞ -manifold in $J^r(M, N)$ in a natural way such that the projections $\pi_s: J^r(M, N) \rightarrow M$ and $\pi_t: J^r(M, N) \rightarrow N$ are both of class C^∞ and $J^r(M, N)$ is a fiber bundle over $M(N)$ with projection $\pi_s(\pi_t)$. As examples, we have:

- (1) $J_0^1(\mathbf{R}, N)$ is identified with the tangent vector bundle of N .
- (2) The set $J^r(\xi)$ of all jets of order r deter-

mined by the sections of a vector bundle ξ of class C^∞ is also a vector bundle of class C^∞ .

Let $f: M \rightarrow N$ and $g: N \rightarrow L$ be C^∞ -mappings. We define a composition of jets by $j_{f(p)}^r g \cdot j_p^r f = j_p^r(g \circ f)$. A jet $j_p^r f \in J^r(M, N)$ is **invertible** if there exists a mapping $g: N \rightarrow M$ such that $j_{f(p)}^r g \cdot j_p^r f = j_p^r(1_M)$, where 1_M denotes the identity mapping of M onto itself. We denote by $I^r(M, N)$ the set of all invertible jets in $J^r(M, N)$, and put $I_p^r(M, N) = I^r(M, N) \cap J_p^r(M, N)$.

(3) Let $G^r(n)$ be the set of all invertible jets in $I^r(\mathbf{R}^n, \mathbf{R}^n)$ whose source and target are the origin of \mathbf{R}^n . Then with respect to the composition of jets, $G^r(n)$ forms a Lie group that is an extension of $G^1(n) = \text{GL}(n, \mathbf{R})$ by a simply connected nilpotent Lie group. The projection $G^r(n) \rightarrow G^1(n)$ is a special example of the natural projection $J^r(M, N) \rightarrow J^s(M, N)$ ($r \geq s$), which is defined in general.

(4) We can identify $I_0^1(\mathbf{R}^m, M)$ ($m = \dim M$) with the tangent m -frame bundle over M . More generally, $I_0^r(\mathbf{R}^m, M)$ is a $G^r(m)$ -bundle over M .

Y. Pseudogroup Structure

Let X be a topological space, and let Γ be a set consisting of homeomorphisms $f: U_f \rightarrow V_f$, where U_f, V_f are open subsets of X . We call Γ a **pseudogroup** of topological transformations if Γ satisfies the following four conditions: (i) Γ contains the identity mapping of X onto X ; (ii) if $f \in \Gamma$, then the restriction of f onto any open subset U of U_f is also contained in Γ ; (iii) if f and g are in Γ and $V_f \subset U_g$, then $g \circ f$ is contained in Γ ; and (iv) if $f \in \Gamma$, then $f^{-1}: V_f \rightarrow U_f$ is also in Γ .

Following the definition of differentiable manifolds we define a **pseudogroup structure** of M (or, more precisely, a Γ -structure of M) as a set A of bijections, with each member α defined on a subset U_α of M onto an open set V_α of X , satisfying the following three conditions: (i) $\bigcup_\alpha U_\alpha = M$; (ii) if $\alpha, \beta \in A$, then $\alpha \circ \beta^{-1} \in \Gamma$, where the domain of definition of $\alpha \circ \beta^{-1}$ is $\beta(U_\alpha \cap U_\beta)$; and (iii) A is the maximal set of bijections that satisfies conditions (i) and (ii). We introduce in M the weakest (coarsest) topology such that every bijection α is a homeomorphism. If Γ' is a pseudogroup of X containing Γ and A' is a Γ' -structure of M such that $A \subset A'$, then we say that A' is subordinate to A . If $X = \mathbf{R}^n$ (or H^n , a half-space of \mathbf{R}^n), Γ is the totality of diffeomorphisms of class C^r of open sets of X onto open sets of X , and M is a space with Hausdorff topology, then the Γ -structure is the C^r -structure with or without boundary which we have already defined. We give three examples of Γ -

structures subordinate to Γ' , where Γ' is the totality of local transformations of class C^r ($r \geq 1$) in \mathbf{R}^n .

(1) When n is even, we identify \mathbf{R}^n with $\mathbf{C}^{n/2}$ and denote the totality of holomorphic transformations of connected open domains by Γ . The Γ -structure in this case is called a **complex structure**.

(2) When n is odd, we define Γ as the totality of transformations of connected open domains in \mathbf{R}^n that leave invariant a Pfaffian form $\sum_{i=1}^m x^i dx^{m+i} + dx^{2m+1}$ ($n = 2m + 1$) up to scalar factors. The Γ -structure in this case is called a **contact structure**.

(3) We consider $\mathbf{R}^n = \mathbf{R}^p \times \mathbf{R}^{n-p}$ and define Γ as the set of all diffeomorphisms $U \rightarrow V$ (where U, V are open in \mathbf{R}^n) such that each set of the form $U \cap (\mathbf{R}^p \times \{y\})$ is mapped onto a set of the form $V \cap (\mathbf{R}^p \times \{y'\})$. The Γ -structure in this case is called a **foliated structure**.

The problem of determining whether there exists a Γ -structure for given Γ and M involves widely ranging problems of topology and analysis. The classification of Γ with reasonable conditions is another important open problem.

Haefliger has constructed the classifying space $B\Gamma$ for Γ -structures.

Z. Infinite-Dimensional Manifolds

Let B and B' be Banach spaces and φ be a mapping from an open set of B to B' . Then, using the notion of Fréchet derivatives, we can define φ to be smooth. For a smooth mapping φ , the Jacobian $J(\varphi)$ at each point x is a linear mapping from B to B' , for which the inverse function theorem holds true as a straightforward extension of the corresponding one in the finite-dimensional case. Similar extension also holds for the existence and the uniqueness theorems of solution of ordinary differential equations with value in B . These facts permit us to generalize the notion of differentiable manifolds to infinite-dimensional ones. Actually, infinite-dimensional manifolds can be defined in a way similar to the finite-dimensional case, taking open sets of a certain Banach space as local coordinate neighborhoods. Such a manifold is called a **Banach manifold**. Various formal definitions of differentiable manifolds can also be stated for Banach manifolds in extended form. However, while differentiable manifolds are locally compact and admit a partition of unity by smooth functions subordinate to a locally finite open covering, Banach manifolds generally lack these properties. Actually, local compactness gives a criterion for whether a manifold is finite-dimensional or not. Banach manifolds

often provide a basic functional-analytic view to nonlinear analysis and global analysis.

The most important category of Banach manifolds is provided by the **Hilbert manifold**, that is, a Banach manifold whose local coordinates are modeled on a Hilbert space. For a Hilbert manifold, a partition of unity subordinate to a locally finite open covering can be taken from smooth functions. In the following, we refer to a separable Hilbert manifold simply as a Hilbert manifold. An infinite-dimensional Hilbert manifold M can be smoothly embedded as an open set of a Hilbert space and thus covered by a single coordinate neighborhood. Hence, in particular, the tangent bundle of M is trivial. Historically, this fact was the first instance showing that the distinguishing properties are shared by infinite-dimensional manifolds; this was first recognized as a consequence of the theorem stating that the unitary group of an infinite-dimensional Hilbert space is contractible. Two infinite-dimensional Hilbert manifolds are diffeomorphic if and only if they are homotopically equivalent. A typical example of a Hilbert manifold is provided by the space of L^2 -loops on a compact Riemannian manifold. Morse theory can be extended in a suitable way to a Riemannian Hilbert manifold under the Palais-Smale condition, which makes it possible for the integral curve of $\text{grad } f$ to tend to a critical point, where f is a Morse function.

AA. Gelfand-Fuks Cohomology

The space $\mathfrak{X}(M)$ consisting of all the smooth vector fields on a smooth manifold M has the structure of a Lie algebra under the bracket operation $[X, Y] = XY - YX$, where the vector fields X and Y are regarded as derivations on the algebra $C^\infty(M)$ of smooth functions on M . $\mathfrak{X}(M)$ is a topological Lie algebra when endowed with the topology defined by uniform convergence of the components of vector fields and all their partial derivatives on each compact set of M . When $\mathfrak{X}(M)$ acts continuously on a topological vector space V , the continuous cohomology $H^*(\mathfrak{X}(M), V)$ of $\mathfrak{X}(M)$ with coefficients in V is the cohomology of the cochain complex $\bigoplus \{C^p = C^p(\mathfrak{X}(M), V), d\}$. Here $C^0 = V$, C^p ($p \geq 1$) is the space of all the alternating p -linear continuous mapping φ of $\mathfrak{X}(M) \times \dots \times \mathfrak{X}(M)$ (p -times) into V , and $d: C^p \rightarrow C^{p+1}$ is defined for $\varphi \in C^p$ and $X_i \in \mathfrak{X}(M)$ by $d\varphi(X_1) = X_1\varphi$ ($p=0$) and $d\varphi(X_1, \dots, X_{p+1}) = \sum_{i < j} (-1)^{i+j} \varphi([X_i, X_j], X_1, \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_{p+1}) + \sum_i (-1)^{i+1} X_i \varphi(X_1, \dots, \hat{X}_i, \dots, X_{p+1})$ ($p \geq 1$). When V is a topological algebra and the elements of $\mathfrak{X}(M)$ act on V as derivations, the exterior multiplication of cochains induces

a graded algebra structure in $H^*(\mathfrak{X}(M), V)$. When $V = \mathbf{R}$ is the trivial $\mathfrak{X}(M)$ -module, then $H^*(\mathfrak{X}(M)) = H^*(\mathfrak{X}(M), \mathbf{R})$ is called the **Gel'fand-Fuks cohomology** of M . Gel'fand and Fuks proved that, for any compact oriented manifold M , we have $\dim H^p(\mathfrak{X}(M)) < +\infty$ for all p and $H^p(\mathfrak{X}(M)) = 0$ for $1 \leq p \leq n$ ($n = \dim M$). For example, if M is the circle S^1 , then the algebra $H^*(\mathfrak{X}(S^1))$ is generated by two generators $\alpha \in H^2$, $\beta \in H^3$, which are explicitly described as cochains in the following way:

$$\alpha\left(f \frac{d}{dt}, g \frac{d}{dt}\right) = \int_{S^1} \begin{vmatrix} f' & f'' \\ g' & g'' \end{vmatrix} dt,$$

$$\beta\left(f \frac{d}{dt}, g \frac{d}{dt}, h \frac{d}{dt}\right) = \int_{S^1} \begin{vmatrix} f & f' & f'' \\ g & g' & g'' \\ h & h' & h'' \end{vmatrix} dt.$$

Localization of the concept of Gel'fand-Fuks cohomology naturally yields the cohomology of **formal vector fields**. Here a formal vector field means the expression $\sum f_\mu(x_1, \dots, x_n) \partial / \partial x_\mu$, f_μ being formal power series in x_1, \dots, x_n . The set of all the formal vector fields forms a Lie algebra \mathfrak{a}_n and the continuous cohomology of \mathfrak{a}_n with respect to the Krull topology is denoted by $H^*(\mathfrak{a}_n)$. Let B_U be the universal classifying space of the unitary group $U(n)$, let $(B_U)_{2n}$ be its $2n$ -skeleton, and let P_{2n} be the canonical principal $U(n)$ -bundle restricted to $(B_U)_{2n}$. Then there is an algebra isomorphism $H^*(\mathfrak{a}_n) \cong H^*(P_{2n}; \mathbf{R})$. This cohomology and its variants play important roles in the theory of foliations.

An important subcomplex $\bigoplus \{C_\lambda^p, d\}$ of $\bigoplus \{C^p, d\}$, the diagonal complex, is defined as $C_\lambda^p = \{\varphi \in C^p \mid \varphi(X_1, \dots, X_p) = 0 \text{ if } \text{supp } X_1 \cap \dots \cap \text{supp } X_p = \emptyset\}$. Here, $\text{supp } X_i$ denotes the support of X_i , that is, $\{x \mid X_i(x) \neq 0\}$. Let P_M be the principal $U(n)$ -bundle associated with the complexified tangent bundle of M . $U(n)$ acts freely on the product $P_M \times P_{2n}$ and the quotient space $E_M = P_M \times P_{2n} / U(n)$ is a fiber bundle over M with fiber P_{2n} . Then, if M is a compact oriented manifold, the cohomology $H_\lambda^p(\mathfrak{X}(M))$ of the diagonal complex is completely determined by the isomorphisms $H_\lambda^p(\mathfrak{X}(M)) \cong H^{p+n}(E_M; \mathbf{R})$ for all p . In particular, if all the Pontryagin classes of M vanish, then $H_\lambda^p(\mathfrak{X}(M)) = \sum_{i+j=p+n} H^i(M; \mathbf{R}) \otimes H^j(\mathfrak{a}_n)$.

The Gel'fand-Fuks cohomology has a topological interpretation: $H^*(\mathfrak{X}(M)) \cong H^*(\Gamma(E_M), \mathbf{R})$ as graded algebras, where $\Gamma(E_M)$ denotes the space of all the continuous cross-sections of $E_M \rightarrow M$ with the compact open topology. Moreover the differential graded algebra $C^*(\mathfrak{X}(M))$ has a model in the sense of Sullivan constructed purely algebraically from a model of the de Rham algebra of M and the Pon-

tryagin classes of M . This model shows in particular that $H^*(\mathfrak{X}(M))$ is not necessarily finitely generated as a graded algebra.

The cohomology theory of $\mathfrak{X}(M)$ in the case where the representation is nontrivial has also been investigated. The natural representation on $C^\infty(M)$ is a typical example. There is also a topological interpretation: $H^*(\mathfrak{X}(M), C^\infty(M)) \cong H^*(Y_M, \mathbf{R})$, where Y_M is the fiber product of the evaluation mapping $M \times \Gamma(E_M) \rightarrow E_M$ and the inclusion $P_M \hookrightarrow E_M$ corresponding to a fiber inclusion $U(n) \hookrightarrow P_{2n}$.

References

- [1] L. Auslander, *Differential geometry*, Harper, 1967.
- [2] R. L. Bishop and S. I. Goldberg, *Tensor analysis on manifolds*, Macmillan, 1968.
- [3] N. Bourbaki, *Variétés différentielles et analytiques*, Hermann, 1967.
- [4] E. Cartan, *Les systèmes différentiels extérieurs et leurs applications géométriques*, *Actualités Sci. Ind.*, Hermann, 1945.
- [5] H. Cartan, *Formes différentielles*, Hermann, 1967.
- [6] C. Chevalley, *Theory of Lie groups I*, Princeton Univ. Press, 1946.
- [7] G. de Rham, *Variétés différentiables*, *Actualités Sci. Ind.*, Hermann, 1955.
- [8] S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, Interscience, I, 1963; II, 1969.
- [9] S. Lang, *Introduction to differentiable manifolds*, Interscience, 1962.
- [10] Y. Matsushima, *Differentiable manifolds*, Dekker, 1972.
- [11] J. R. Munkres, *Elementary differential topology*, *Ann. Math. Studies* 54, Princeton Univ. Press, 1963.
- [12] K. Nomizu, *Lie groups and differential geometry*, *Publ. Math. Soc. Japan*, 1956.
- [13] S. Sternberg, *Lectures on differential geometry*, Prentice-Hall, 1964.
- [14] H. Whitney, *Differentiable manifolds*, *Ann. Math.*, (2) 37 (1936), 645–680.
- [15] H. Whitney, *Geometric integration theory*, Princeton Univ. Press, 1957.
- [16] J. Eells, A setting for global analysis, *Bull. Amer. Math. Soc.*, 72 (1966), 751–807.
- [17] D. Burghlelea and N. Kuiper, *Hilbert manifolds*, *Ann. Math.*, (2) 90 (1969), 335–352.
- [18] N. Kuiper, The homotopy type of the unitary group of Hilbert space, *Topology*, 3 (1965), 19–30.
- [19] R. Palais and S. Smale, A generalized Morse theory, *Bull. Amer. Math. Soc.*, 70 (1964), 156–172.
- [20] I. M. Gel'fand, The cohomology of infinite-dimensional Lie algebras; some ques-

tions of integral geometry, Actes Congr. Int. Math., Nice, 1970, Gauthier-Villars, vol. 1, 95–111.

[21] I. M. Gel'fand and D. B. Fuks, Cohomologies of the Lie algebra of tangent vector fields of a smooth manifold I, II, Functional Anal. Appl., 3 (1969), 194–210; 4 (1970), 110–116. (Original in Russian, 1969, 1970.)

[22] I. M. Gel'fand and D. B. Fuks, The cohomology of the Lie algebra of formal vector fields, Math. USSR-Izv. (1970), 322–337. (Original in Russian, 1970.)

[23] V. W. Guillemin, Cohomology of vector fields on a manifold, Advances in Math., 10 (1973), 192–220.

[24] M. V. Losik, On the cohomologies of infinite-dimensional Lie algebras of vector fields, Functional Anal. Appl., 4 (1970), 127–135. (Original in Russian, 1970.)

[25] R. Bott and G. Segal, The cohomology of the vector fields on a manifold, Topology, 16 (1977), 285–298.

[26] A. Haeffliger, Sur la cohomologie de l'algèbre de Lie des champs de vecteurs, Ann. Sci. Ecole Norm. Sup., 9 (1976), 503–532.

106 (X.6) Differential Calculus

A. First-Order Derivatives

Let $y = f(x)$ be a real-valued function of x defined on an interval I of the real line \mathbf{R} . If for a fixed $x_0 \in I$, the limit

$$\lim_{\substack{h \rightarrow 0 \\ x_0 + h \in I}} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists and is finite, then f is called **differentiable at the point** x_0 , and the limit is the **derivative (differential coefficient or differential quotient)** of f at the point x_0 . If f is differentiable at every point of a set $A \subset I$, then f is said to be **differentiable on** A . The function that assigns the derivative of f at x to $x \in A$ is called the **derivative (or derived function)** of $f(x)$, which is denoted by dy/dx , y' , \dot{y} , $df(x)/dx$, $(d/dx)f(x)$, $f'(x)$, or $D_x f(x)$. The process of determining f' is known as the **differentiation** of f . The derivative of f at the point x_0 is written $f'(x_0)$, $(df/dx)(x_0)$, $D_x f(x_0)$, $[dy/dx]_{x=x_0}$, etc. We say that f is **right (left) differentiable** or **differentiable on the right (left)** if the limit on the right, $\lim_{h \rightarrow +0} (f(x_0 + h) - f(x_0))/h$ (the limit on the left, $\lim_{h \rightarrow +0} (f(x_0 - h) - f(x_0))/h$), exists and is finite. This limit is called the **right (left) derivative** or **derivative on the right (left)** and is

denoted by $D_x^+ f(x_0)$ or $f'_+(x_0)$ ($D_x^- f(x_0)$ or $f'_-(x_0)$). For instance, if f is defined on $I = [a, b]$, then $D_x f(a)$ is identical to $D_x^+ f(a)$.

B. Differentials

In the definitions given above, neither dx nor dy in dy/dx has a meaning by itself. In the following, however, we give a definition of dx and dy , using the concept of increment, so that $dy = f'(x)dx$. Let Δy denote the **increment** $f(x + \Delta x) - f(x)$ of f corresponding to the increment Δx of x . Suppose that $f(x)$ is differentiable at x . We set $\Delta y/\Delta x = f'(x) + \varepsilon$. Then we have $\lim_{\Delta x \rightarrow 0} \varepsilon = 0$. This can be written utilizing Landau's notation as $\Delta y = f'(x)\Delta x + o(|\Delta x|)$ ($\Delta x \rightarrow 0$); in other words, Δy is the sum of two terms, of which the first, $f'(x)\Delta x$, is proportional to Δx and the second is an infinitesimal of an order higher than Δx . Here the principal part $f'(x)\Delta x$ of Δy is called the **differential** of $y = f(x)$ and is denoted by dy . The differential dy thus defined is a function of two independent variables x and Δx . In particular, if $f(x) = x$, from the definition we get $dx = 1 \cdot \Delta x = \Delta x$. Hence, in general, we have $dy = f'(x)dx$ and $f'(x) = dy/dx$.

With respect to the rectangular coordinates (x, y) , the straight line with slope $f'(x_0)$ through a point $(x_0, f(x_0))$ on the graph of $y = f(x)$ is the **tangent line** of the graph at the point $(x_0, f(x_0))$. A function is continuous at a point where the function is differentiable, but the converse of this proposition does not hold. In fact, Weierstrass showed that the function defined by the infinite series $\sum_{n=0}^{\infty} a^n \cos b^n \pi x$, where $0 < a < 1$ and b is an odd integer with $ab > (3/2)\pi + 1$, is continuous everywhere and nowhere differentiable on $(-\infty, \infty)$ [3].

C. Differentiation

For two differentiable functions f and g defined on the interval I , the following formulas hold: $(\alpha f + \beta g)' = \alpha f' + \beta g'$, where α and β are constants; $(fg)' = f'g + fg'$; and $(f/g)' = (f'g - fg')/g^2$ (at every point where $g \neq 0$). Let $y = f(x)$ be a function of x defined on the interval (a, b) and $x = \varphi(t)$ a function of t defined on (α, β) . If $\varphi(t) \in (a, b)$ whenever $t \in (\alpha, \beta)$, then the composite function $y = F(t) = f(\varphi(t)) = (f \circ \varphi)(t)$ is well defined. Assume further that f and φ are differentiable on (a, b) and (α, β) , respectively. Then the composite function $F(t) = (f \circ \varphi)(t)$ is differentiable on (α, β) , and we have the **chain rule**, $F'(t) = f'(x)\varphi'(t)$ ($x = \varphi(t)$), or $dy/dt = (dy/dx)(dx/dt)$. Assume that a function $y = f(x)$ is **strictly increasing** or **decreasing** and differentiable at x_0 . If furthermore we have

$f'(x) \neq 0$, then the inverse function $x = f^{-1}(y)$ is also differentiable at $y_0 (= f(x_0))$ and satisfies $(dx/dy)_{y=y_0} (dy/dx)_{x=x_0} = 1$. However, if $f'(x_0) = 0$, then even though $f^{-1}(y)$ is not differentiable at y_0 , $\lim_{\Delta y \rightarrow 0} (f^{-1}(y_0 + \Delta y) - f^{-1}(y_0))/\Delta y$ exists and is $+\infty$ or $-\infty$.

D. Higher-Order Derivatives

If the derivative $f'(x)$ of a function $y = f(x)$ is again differentiable on I , then $(f'(x))' = f''(x)$ is well defined as a function of x on I . In general, if $f^{(n-1)}(x)$ is differentiable on I , then $f(x)$ is called **n -times differentiable** on I , and the **n th derivative** (or **n th derived function**) $f^{(n)}(x)$ of $f(x)$ is defined by $f^{(n)}(x) = (f^{(n-1)}(x))'$ and is also denoted by $d^n y/dx^n$ or $D^{(n)}y$. The n th derivative for $n \geq 2$ is called a **higher-order derivative**.

Concerning the n th derivative of the product of two functions, **Leibniz's formula** holds:

$$(fg)^{(n)} = f^{(n)}g + \binom{n}{1}f^{(n-1)}g' + \dots + \binom{n}{k}f^{(n-k)}g^{(k)} + \dots + fg^{(n)}.$$

Analogous to $dy = y'\Delta x$, which is a function of x and Δx , we can define d^2y in the notation d^2y/dx^2 by $d^2y = d(dy) = d(y'\Delta x) = (y'\Delta x)'\Delta x = y''\Delta x^2$. Since $\Delta x = dx$, it follows from the above that $d^2y = y''dx^2$. Similarly, $d^n y = y^{(n)}dx^n$ and is called the **n th differential** (or **differential of n th order**) of $f(x)$.

E. The Mean Value Theorem

Let $f(x)$ be a continuous function defined on $[a, b]$, and suppose that for every point x_0 on (a, b) there exists a limit $\lim_{h \rightarrow 0} (f(x_0 + h) - f(x_0))/h$, which may be infinite. (These conditions are satisfied if $f(x)$ is differentiable on $[a, b]$.) Then there exists a point ξ such that

$$\frac{f(b) - f(a)}{b - a} = f'(\xi), \quad a < \xi < b.$$

This proposition is called the **mean value theorem**. A special case of the theorem under the further condition that $f(a) = f(b)$ is called **Rolle's theorem**. If we put $b - a = h$, $\xi = a + \theta h$, then the conclusion of the theorem may be written as $f(a + h) = f(a) + h \cdot f'(a + \theta h)$ ($0 < \theta < 1$).

This theorem implies the following: Let $f(x)$ be a function as in the hypothesis of the mean value theorem, and assume further that $A \leq f'(x) \leq B$ holds for all x with $a < x < b$. Then $A \leq (f(b) - f(a))/(b - a) \leq B$. (French mathematicians sometimes call this the "thé-

ème des accroissements finis.") Using the mean value theorem, the following theorem can be proved: If $f(x)$ is continuous on $[a, b]$ and $f'(x)$ exists and is positive on (a, b) , then $f(b) > f(a)$. Accordingly, if $f'(x) > 0$ at every point x of an interval I , then $f(x)$ is **strictly increasing** on that interval. (If $f'(x) < 0$ on I , then f is **strictly decreasing**.) The converse of the previous statement does not always hold ($f'(x) = x^3$ is a counterexample, since $f'(0) = 0$). Furthermore, from the mean value theorem it follows that if $f'(x) = 0$ everywhere in an interval, then $f(x)$ is constant on that interval. Consequently, two functions with the same derivative on an interval differ only by a constant.

Suppose that $f(x)$ is n -times differentiable on an open interval I . For a fixed $a \in I$ and an arbitrary $x \in I$, we put

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \dots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + R_n.$$

Then $R_n = f^{(n)}(\xi)(x-a)^n/n!$ for some ξ between a and x . This is called **Taylor's formula**, where R_n is the **remainder** of the n th order given by Lagrange. We also have several other forms for R_n (\rightarrow Appendix A, Table 9). If $f^{(n)}(x)$ is continuous at $x = a$, then $\xi \rightarrow a$ as $x \rightarrow a$, and accordingly, $f^{(n)}(\xi) \rightarrow f^{(n)}(a)$. Hence $f(x) = \sum_{k=0}^n (f^{(k)}(a)/k!)(x-a)^k + o((x-a)^n)$. If $f^{(n)}(x)$ is continuous at $x = a$, then, by Taylor's formula, the value of the polynomial $\sum_{k=0}^n (f^{(k)}(a)/k!)(x-a)^k$ can be considered an approximate value of $f(x)$ for x near a . This approximation is called the **n th approximation** of $f(x)$, and its error is given by $|R_{n+1}|$. By applying this formula, it is sometimes possible to calculate a limit such as $A = \lim_{x \rightarrow a} f(x)/g(x)$, where $f(x) \rightarrow 0$ and $g(x) \rightarrow 0$ as $x \rightarrow a$. For instance, if $f'(x)$ and $g'(x)$ are both continuous at $x = a$ and $g'(a) \neq 0$, then by taking the first approximations of $f(x)$ and $g(x)$ it is easily seen that $A = f'(a)/g'(a)$. A limit of this type is often called a **limit of an indeterminate form** $0/0$. Similarly, we can calculate limits of such indeterminate forms as $0 \cdot \infty$ or 0^∞ (for limits of indeterminate forms \rightarrow [6]).

F. Partial Derivatives

Let $w = f(x, y, \dots, z)$ be a real-valued function of n independent real variables x, y, \dots, z defined on a domain G contained in n -dimensional Euclidean space \mathbf{R}^n . We obtain a function of a single variable from f by keeping $n-1$ variables (say, (x, y, \dots, z) , i.e., all except y) fixed. If

such a function $\varphi(y) = f(x_0, y, \dots, z_0)$ is differentiable, that is, if

$$\begin{aligned}\varphi'(y_0) &= \lim_{\Delta y \rightarrow 0} \frac{\varphi(y_0 + \Delta y) - \varphi(y_0)}{\Delta y} \\ &= \lim_{\Delta y \rightarrow 0} \frac{f(x_0, y_0 + \Delta y, \dots, z_0) - f(x_0, y_0, \dots, z_0)}{\Delta y}\end{aligned}$$

exists and is finite, then f is called **partially differentiable** with respect to y at (x_0, y_0, \dots, z_0) , and the derivative is called the **partial derivative** (or **partial differential coefficient**) of $f(x, y, \dots, z)$ with respect to y at (x_0, y_0, \dots, z_0) . It is denoted by $[\partial w / \partial y]_{x=x_0, \dots, z=z_0}$, $(\partial / \partial y)f(x_0, y_0, \dots, z_0)$, $f_y(x_0, y_0, \dots, z_0)$, or $D_y f(x_0, y_0, \dots, z_0)$, etc. We usually assume that the point (x, y, \dots, z) where partial derivatives are considered is an interior point of the domain of the function. Since in a space of dimension higher than 1, the boundary of a domain may be complicated, partial derivatives at boundary points are usually not considered. If a function f possesses a partial derivative with respect to x at every point of an open set G , then f_x is a function on G and is called a **partial derivative** of f with respect to x . The process of determining partial derivatives of f is called the **partial differentiation** of f .

G. Total Differential

Let $w = f(x, y, \dots, z)$ be a function defined on a domain G , and let $P = (x, y, \dots, z)$ be an interior point of the domain G of a function $w = f(x, y, \dots, z)$. Put $\Delta w = f(x_0 + \Delta x, y_0 + \Delta y, \dots, z_0 + \Delta z) - f(x_0, y_0, \dots, z_0)$. If there exist constants $\alpha, \beta, \dots, \gamma$ such that $\Delta w = \alpha \Delta x + \beta \Delta y + \dots + \gamma \Delta z + o(\rho)$ ($\rho \rightarrow 0$), where $\rho = \sqrt{\Delta x^2 + \Delta y^2 + \dots + \Delta z^2}$, then f is called **totally differentiable** (or **differentiable in the sense of Stolz**) at P . In this case, f is partially differentiable at P with respect to each of the variables x, y, \dots, z , and $\alpha = f_x(x_0, y_0, \dots, z_0)$, $\beta = f_y(x_0, y_0, \dots, z_0)$, \dots , $\gamma = f_z(x_0, y_0, \dots, z_0)$. The principal part of Δw as $\rho \rightarrow 0$ is $\alpha \Delta x + \beta \Delta y + \dots + \gamma \Delta z$, which is called the **total differential** of w at P . If f is totally differentiable at every point of G , then f is said to be totally differentiable on G . The total differential of w is denoted by dw . Since the total differentials of x, y, \dots, z are $dx = \Delta x, dy = \Delta y, \dots, dz = \Delta z$, respectively, we can write $dw = f_x dx + f_y dy + \dots + f_z dz$; and dw is a function of independent variables $x, y, \dots, z, dx, dy, \dots, dz$. The total differentiability of f implies the continuity of f , whereas the partial differentiability of f with respect to each variable does not imply that f is continuous. (Example: Define $f(x, y) = xy/(x^2 + y^2)$ for $(x, y) \neq (0, 0)$, and $f(0, 0) = 0$; then the func-

tion f is not continuous at $(0, 0)$, even though both f_x and f_y exist at $(0, 0)$.) The function f is totally differentiable on G if all f_x, f_y, \dots, f_z exist and are continuous on G , or, more weakly, if all f_x, f_y, \dots, f_z exist and, with possibly one exception, are continuous. Suppose that $w = f(x, y)$ is totally differentiable at (x, y) , and let $\Delta x = \rho \cos \theta, \Delta y = \rho \sin \theta$. As $\rho \rightarrow 0$, for a fixed θ there exists the limit $\lim_{\rho \rightarrow 0} (\Delta w / \rho) = f_x(x, y) \cos \theta + f_y(x, y) \sin \theta$. This limit is called the **directional derivative** in the direction θ at (x, y) . The partial derivatives f_x and f_y are special cases of the directional derivative for $\theta = 0$ and $\pi/2$, respectively. Suppose that we are given a curve lying in the interior of the domain of f and that the curve passes through the point (x, y) , where the curve is differentiable. Then the partial derivative of $w = f(x, y)$ in the direction of the normal to the curve at (x, y) is called the **normal derivative** of w at the point (x, y) on the curve and is denoted by $\partial w / \partial n$. Analogous definitions and notations have been introduced for functions of more than two variables.

To see the geometric significance of the total differentiability of $w = f(x, y)$, we consider the graph of the function $w = f(x, y)$ and a point $(a, b, f(a, b))$ on the graph. Then the plane represented by $w - f(a, b) = \alpha(x - a) + \beta(y - b)$ is the tangent plane to the surface at $(a, b, f(a, b))$ if and only if $\alpha = f_x(a, b)$ and $\beta = f_y(a, b)$. The existence of f_x and f_y depends on the choice of coordinate axes, while the total differentiability of f does not.

H. Higher-Order Partial Derivatives

Suppose that a partial derivative of a function $w = f(x, y, \dots, z)$ defined on an open set G again admits partial differentiation. The latter partial derivative is called a second-order partial derivative of f . We can similarly define the n th order partial derivatives. Higher-order partial derivatives are denoted as follows:

$$\frac{\partial}{\partial x} \left(\frac{\partial w}{\partial x} \right) = \frac{\partial^2 w}{\partial x^2} = f_{xx}(x, y, \dots, z),$$

$$\frac{\partial}{\partial y} \left(\frac{\partial w}{\partial x} \right) = \frac{\partial^2 w}{\partial x \partial y} = f_{xy}(x, y, \dots, z),$$

$$\frac{\partial}{\partial x} \left(\frac{\partial^2 w}{\partial x \partial y} \right) = \frac{\partial^3 w}{\partial x \partial y \partial x} = f_{xyx}(x, y, \dots, z), \dots$$

In general, f_{xy} and f_{yx} are not equal. (Peano's example: Let $f(x, y) = xy(x^2 - y^2)/(x^2 + y^2)$ for $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$. Then $f_{xy} = -1$, $f_{yx} = 1$ at $(0, 0)$.) However, if both f_{xy} and f_{yx} are continuous on an open set G' , then they coincide in G' . Furthermore, if f_x, f_y , and f_{xy} exist in a neighborhood U of a point P belonging to the domain of f and f_{xy} is continuous

at P , then f_{yx} exists at P and $f_{xy} = f_{yx}$ (H. A. Schwarz). If f_x and f_y exist in U and are totally differentiable at P , then $f_{xy} = f_{yx}$ at P (W. H. Young). Similarly, if the partial derivatives of order ≥ 3 $f_{\dots xy \dots}$ and $f_{\dots yx \dots}$ are all continuous, then $f_{\dots xy \dots} = f_{\dots yx \dots}$. Hence we can change the order of differentiation if all the derivatives concerned are continuous.

I. Composite Functions of Several Variables

Let w be a function of x, y, \dots, z , and let each x, y, \dots, z be a function of t . Suppose that the range of $(x(t), y(t), \dots, z(t))$ is contained in the domain of w . Then w is a function of t . If further w is totally differentiable and x, y, \dots, z are all differentiable, then w as a function of t is differentiable, and we have

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \dots + \frac{\partial w}{\partial z} \frac{dz}{dt}.$$

If partial derivatives of order ≥ 2 are totally differentiable, then $d^2w/dt^2, d^3w/dt^3, \dots$ are obtained by repeating the above procedure. A similar consideration is valid when x, y, \dots, z are functions of several variables.

J. Taylor's Formula for Functions of Several Variables

Suppose that $f(x, y)$ is defined on an open domain G , $f(x, y)$ has continuous partial derivatives of orders up to n , and the line segment $(a + (x-a)t, b + (y-b)t)$ ($0 \leq t \leq 1$) is contained in the domain G . Then there exists a number θ ($0 < \theta < 1$) such that

$$\begin{aligned} f(x, y) &= f(a, b) + \left((x-a) \frac{\partial}{\partial x} + (y-b) \frac{\partial}{\partial y} \right) f(a, b) \\ &+ \frac{1}{2!} \left((x-a) \frac{\partial}{\partial x} + (y-b) \frac{\partial}{\partial y} \right)^2 f(a, b) + \dots \\ &+ \frac{1}{(n-1)!} \left((x-a) \frac{\partial}{\partial x} + (y-b) \frac{\partial}{\partial y} \right)^{n-1} f(a, b) \\ &+ \frac{1}{n!} \left((x-a) \frac{\partial}{\partial x} + (y-b) \frac{\partial}{\partial y} \right)^n \\ &\quad \times f(a + (x-a)\theta, b + (y-b)\theta), \end{aligned}$$

where, for instance, the third term $((x-a) \cdot (\partial/\partial x) + (y-b)(\partial/\partial y))^2 f(a, b)$ means $(x-a)^2 (\partial^2 f / \partial x^2)(a, b) + 2(x-a)(y-b) \cdot (\partial^2 f / \partial x \partial y)(a, b) + (y-b)^2 (\partial^2 f / \partial y^2)(a, b)$, with $(\partial^2 f / \partial x^2)(a, b)$, $(\partial^2 f / \partial x \partial y)(a, b)$, and $(\partial^2 f / \partial y^2)(a, b)$ denoting the values of $(\partial^2 f / \partial x^2)$, $(\partial^2 f / \partial x \partial y)$, and $(\partial^2 f / \partial y^2)$ at (a, b) , respectively. The displayed formula is called **Taylor's formula** for a function of two variables. A similar

formula is valid for a function of n variables ($n \geq 3$). As in the case of functions of one variable, we can derive approximation formulas for f from Taylor's formula.

K. Classes of Functions

If all the partial derivatives of order n of $f(P)$ are continuous on an open set G , then f is said to be a **function of class C^n** (or **n -times continuously differentiable**) on G . The set of all n -times continuously differentiable functions is denoted by C^n ($n = 1, 2, \dots$). A continuous function is of class C^0 . A **function of class C^1** is also called a **smooth function**. It is obvious that $C^0 \supset C^1 \supset C^2 \supset \dots$. A partial derivative of order $r \leq s$ of a function belonging to class C^s does not depend on the order of the differentiation. A function belonging to $C^\infty = \bigcap_{r=1}^\infty C^r$ is said to be of class C^∞ or **infinitely differentiable**. We sometimes say that a function has a certain "nice" property or is "well behaved" if it belongs to some C^r ($r \geq 1$).

Let $w = f(x, y, \dots, z)$ be a function defined on an open set G in \mathbf{R}^n and $P = (a, b, \dots, c) \in G$. If

$$f(x, y, \dots, z) = f(a, b, \dots, c) + \sum_{r_1=1}^\infty \dots \sum_{r_n=1}^\infty \alpha_{r_1 r_2 \dots r_n} (x-a)^{r_1} (y-b)^{r_2} \dots (z-c)^{r_n}$$

holds in some open neighborhood U of P , where the right-hand side of the equality is an absolutely convergent series, then f is said to be **real analytic** at P . In this case, f is r -times differentiable at P for any r , and we have

$$\begin{aligned} \alpha_{r_1 r_2 \dots r_n} &= \frac{r_1! r_2! \dots r_n!}{(r_1 + r_2 + \dots + r_n)!} \\ &\quad \times \frac{\partial^{r_1+r_2+\dots+r_n} f}{\partial x^{r_1} \partial y^{r_2} \dots \partial z^{r_n}}(a, b, \dots, c). \end{aligned}$$

If f is real analytic at every point P of the domain G , then f is called a **real analytic function** on G . Sometimes, a real analytic function is called a function of class C^∞ . A real analytic function belongs to C^∞ , but the converse is not true (\rightarrow 58 C^∞ -Functions and Quasi-Analytic Functions E).

L. Extrema

Let f be a real-valued function defined on a domain G in an n -dimensional Euclidean space \mathbf{R}^n that has the point P_0 in its interior. If there exists a neighborhood U of P_0 such that for every point P ($\neq P_0$) of U we have $f(P) \geq f(P_0)$, then we say that f has a **relative minimum** at P_0 , and $f(P_0)$ is a **relative minimum** of f . Replacing \geq by \leq , we obtain the definition of a **relative maximum**. $f(P_0)$ is

called a **relative extremum** if it is either a relative maximum or a relative minimum.

To find relative extrema, the following facts concerning the sign of the derivative are useful. Suppose that a function of a single variable is differentiable on an interval I . Then we have the following: (1) If f has a relative extremum at an interior point x_0 of I , then $f'(x_0) = 0$. (2) If $f'(x_0) = 0$ and $f'(x)$ changes its sign at x_0 from positive (negative) to negative (positive), then f has a relative maximum (minimum) at x_0 . (3) If $f'(x_0) = 0$ and f is twice differentiable on some neighborhood of x_0 , then f has a relative maximum or minimum according as $f''(x_0) < 0$ or > 0 . If $f''(x_0) = 0$, then nothing definite can be concluded about a relative extremum of f at x_0 . In general, if there exists a neighborhood of x_0 in which f is r -times differentiable (r is even) and $f^{(r)}$ is continuous, and if $f'(x_0) = f''(x_0) = \dots = f^{(r-1)}(x_0) = 0$, $f^{(r)}(x_0) > 0$ (or < 0), then f has a relative minimum (maximum) at x_0 . On the other hand, if this condition holds with odd r , then $f(x_0)$ is not a relative extremum. If $f'(x_0) = 0$, then $f(x_0)$ is called a **stationary value** of f .

If a function f on n variables x, y, \dots, z has a relative extremum at (x_0, y_0, \dots, z_0) , then we have $f_x(x_0, y_0, \dots, z_0) = 0$, $f_y(x_0, y_0, \dots, z_0) = 0$, \dots , $f_z(x_0, y_0, \dots, z_0) = 0$, provided that the partial derivatives of f exist. Assume that for a function f of class C^2 of two variables x and y , we have $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$, and let $\delta = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0)$. Then we have the following: (1) If $\delta > 0$, then according as $f_{xx}(x_0, y_0) < 0$ or > 0 , f has a relative maximum or minimum at (x_0, y_0) . (2) If $\delta < 0$, then f does not have a relative extremum at (x_0, y_0) . (3) If $\delta = 0$, then without further information nothing definite can be said about a relative extremum of f at the point.

Let x_1, \dots, x_n be independent variables. If a function f of variables x_1, \dots, x_n has a relative extremum at a point $P_0 = (x_1^0, \dots, x_n^0)$, then $f_i = f_{x_i}(P_0) = 0$ ($i = 1, \dots, n$), provided that all the partial derivatives of f exist. In general, a point P_0 where f is totally differentiable and these conditions are satisfied is called a **critical point** of f . The value $f(P_0)$ at a critical point is called a **stationary value**. If further f is of class C^2 , then consider a *quadratic form of n variables $Q = Q(X_1, \dots, X_n) = \sum_{i,k} f_{i,k} X_i X_k$, where $f_{i,k} = f_{x_i x_k}(P_0)$. Suppose that $|f_{i,k}| \neq 0$. Then according to whether Q is *positive definite, *negative definite, or *indefinite, f has a relative minimum, relative maximum, or no relative extremum at P_0 . If $|f_{i,k}| = 0$, then nothing can be said in general. A critical point P of f is said to be **nondegenerate** if $|f_{i,k}| \neq 0$ and **degenerate** if $|f_{i,k}| = 0$.

We can also apply the method of differenti-

ation of *implicit functions to find relative extrema of functions defined implicitly. Given functions $\varphi_1, \dots, \varphi_m$ ($m < n$), the problem of finding a relative extremum of $f(x_1, \dots, x_n)$ under the condition that $\varphi_1(x_1, \dots, x_n) = 0, \dots, \varphi_m(x_1, \dots, x_n) = 0$ is called the problem of finding a **conditional relative extremum**.

This problem can be reduced to the problem of finding a relative extremum of an implicit function. Actually, if the functions $f, \varphi_1, \dots, \varphi_m$ are of class C^1 and the *Jacobian $\partial(\varphi_1, \dots, \varphi_m)/\partial(x_{n-m+1}, \dots, x_n)$ does not vanish in the domain considered, then $y_1 = x_{n-m+1}, \dots, y_m = x_n$ can be regarded as implicit functions of x_1, \dots, x_l ($l = n - m$). Hence we can set $f(x_1, \dots, x_l, y_1, \dots, y_m) = f^*(x_1, \dots, x_l)$. Then f has a relative extremum at (x_1^0, \dots, x_n^0) under the condition $\varphi_1 = \dots = \varphi_m = 0$ if and only if f^* has a relative extremum at $P_0 = (x_1^0, \dots, x_l^0)$. The latter condition implies that all $\partial f^*/\partial x_j$ ($j = 1, \dots, l$) vanish at P_0 , which holds if and only if for arbitrary constants $\lambda_1, \dots, \lambda_m$ the function $F(x_1, \dots, x_n) = f + \lambda_1 \varphi_1 + \dots + \lambda_m \varphi_m$ satisfies $\partial F/\partial x_i = 0$ ($i = 1, \dots, n$), and further $\varphi_1 = 0, \dots, \varphi_m = 0$ at (x_1^0, \dots, x_n^0) . From this system of equations we can often find the values of x_1^0, \dots, x_n^0 . This method of finding conditional relative extrema is called **Lagrange's method of indeterminate coefficients** or the **method of Lagrange multipliers** (\rightarrow 208 Implicit Functions; 216 Integral Calculus H; 379 Series H).

References

- [1] T. M. Apostol, *Mathematical analysis*, Addison-Wesley, 1957.
- [2] N. Bourbaki, *Eléments de mathématique, Fonctions d'une variable réelle*, Actualités Sci. Ind., 1074b, 1132a, Hermann, second edition, 1958, 1961.
- [3] R. C. Buck, *Advanced calculus*, McGraw-Hill, second edition, 1965.
- [4] R. Courant, *Differential and integral calculus I, II*, Nordemann, 1938.
- [5] G. H. Hardy, *A course of pure mathematics*, Cambridge Univ. Press, seventh edition, 1938.
- [6] E. Hille, *Analysis*, Blaisdell, I, 1964; II, 1966.
- [7] W. Kaplan, *Advanced calculus*, Addison-Wesley, 1952.
- [8] E. G. H. Landau, *Einführung in die Differentialrechnung und Integralrechnung*, Noordhoff, 1934; English translation, *Differential and integral calculus*, Chelsea, 1965.
- [9] J. M. H. Olmsted, *Advanced calculus*, Appleton-Century-Crofts, 1961.
- [10] A. Ostrowski, *Vorlesungen über*

- Differential-und Integralrechnung I, II, III, Birkhäuser, second edition, 1960–1961.
- [11] M. H. Protter and C. B. Morrey, *Modern mathematical analysis*, Addison-Wesley, 1964.
- [12] W. Rudin, *Principles of mathematical analysis*, McGraw-Hill, second edition, 1964.
- [13] V. I. Smirnov, *A course of higher mathematics. I, Elementary calculus; II, Advanced calculus*, Addison-Wesley, 1964.
- [14] A. E. Taylor, *Advanced calculus*, Ginn, 1955.

107 (XIII.1) Differential Equations

A. Ordinary Differential Equations

It was Galileo who found that the acceleration of a falling body is a constant and thence derived his law of a falling body $x(t) = gt^2/2$ as what we would now view as a solution of the differential equation $x''(t) = g$, where $x(t)$ denotes the distance the body has fallen during the time interval t and g is the constant gravitational acceleration. This pioneering work may be regarded as the first example of solution of a differential equation. Also, the \dagger equations of motion, proposed by I. \dagger Newton as the mathematical formulation of the law of motion, including Galileo's law as a special case, are differential equations of the second order. Thus differential equations appeared, simultaneously with differential and integral calculus, as an indispensable tool for the unified and concise expression of the laws of nature. Such laws are generally called **differential laws**.

Newton completely solved the equations of the \dagger two-body problem proposed by himself; G. W. \dagger Leibniz also succeeded in solving many simple differential equations.

In the 18th century, many mathematicians, such as the \dagger Bernoullis, A. C. Clairaut, J. F. Riccati, L. \dagger Euler, and J. L. \dagger Lagrange, attacked and solved differential equations of various types independently. In that period, the emphasis was on solution by **quadrature**, that is, applying to \dagger elementary functions a finite number of algebraic operations, transformations of variables, and indefinite integrations. It was toward the end of the 18th century that new methods, such as integration by infinite series, came to be discussed. A method of variation of constants for the solution of linear ordinary differential equations was invented by Lagrange in 1775. At the beginning of the 19th century, C. F. \dagger Gauss

initiated the study of differential equations satisfied by \dagger hypergeometric series.

The problem of existence of solutions, which supplies a foundation of modern differential equation theory, was first treated by A. L. \dagger Cauchy. His proof of the existence theorem was later improved by R. L. Lipschitz (1869).

Pioneers in the function-theoretic treatment of differential equations were C. A. A. Briot and J. C. Bouquet, who investigated the singular points of a function defined by an analytic differential equation. Also, B. \dagger Riemann proposed a new viewpoint which influenced L. Fuchs in his development of the theory of linear ordinary differential equations in the complex domain (1865). Works of A. M. Legendre on \dagger elliptic functions and of H. \dagger Poincaré on \dagger automorphic functions should also be mentioned in this connection.

After the Cauchy-Lipschitz existence theorem for the equation $y' = f(x, y)$ was known, efforts were directed toward weakening the conditions imposed on $f(x, y)$. G. Peano first succeeded in giving a proof under the continuity assumption only (1890), and his results were sharpened by O. Perron (1915).

Regarding the uniqueness of solutions of \dagger initial value problems, there are various results by W. F. Osgood (1898), Perron (1925), and many Japanese mathematicians. In the course of this work the necessary and sufficient condition for uniqueness was successfully formulated in a concise form (\rightarrow 316 Ordinary Differential Equations (Initial Value Problems)).

For linear differential equations with periodic coefficients, investigations were carried out by C. Hermite (1877), E. Picard (1881), G. Floquet (1883), G. W. Hill (1886), and others. For instance, solutions satisfying $y(x + \omega) = \lambda y(x)$ were found to exist, where ω is the period of the coefficients. Analogous results followed in the case of doubly periodic coefficients.

Techniques of factorization of linear differential equations developed by G. Frobenius (1873) and E. Landau (1920) should also be noted. Picard (1883), J. Drach (1898), and E. Vessiot (1903, 1904) established a remarkable result on the solvability (in the sense of solution by quadrature) of linear differential equations, successfully extending the \dagger Galois theory in this new direction.

The concept of \dagger asymptotic series, which in a sense approximate the solution of differential equations, was introduced by Poincaré (1886) and extended by M. A. Lyapunov (1892), J. C. C. Kneser (1896), J. Horn (1897), C. E. Love (1914), and others. Poincaré was also the founder of topological methods in differential

equation theory, and his ideas were developed extensively by I. Bendixson (1900), Perron (1922, 1923), G. D. Birkhoff, and others (→ 126 Dynamical Systems).

In 1890, Picard invented an ingenious technique of [†]successive approximation for the proof of existence theorems, and his technique is now widely used in every application of functional equations. The technique of reducing linear differential equations to linear [†]integral equations of Volterra type was also developed.

On the [†]boundary value problems and [†]eigenvalue problems that appear in many areas of physics, there was extensive research by mathematicians such as J. C. F. Sturm (1836), J. Liouville, L. Tonelli, Picard, M. Bôcher (1898, 1921), Birkhoff (1901, 1911), and others. In this connection the problem arises of expanding a given function by an [†]orthogonal system of functions obtained as [†]eigenfunctions of a given boundary value problem. Those problems were brought into unified form by D. [†]Hilbert (1904) in his theory of [†]integral equations. Subsequently boundary value problems of ordinary and partial differential equations came to be discussed in this framework.

Finally, it should be mentioned that the [†]calculus of variations created by Euler and Lagrange gave rise to the study of a certain class of differential equations bearing the name of Euler (→ 46 Calculus of Variations).

B. Partial Differential Equations

The origin of partial differential equations can be traced back to the study of hydrodynamic problems by J. d'Alembert (1744) and Euler. However, perhaps Lagrange and P. S. [†]Laplace were the first to investigate the general theory. Subsequently, during the 18th and 19th centuries, it was developed by G. Monge, A.-M. Ampère, J. F. Pfaff, C. G. [†]Jacobi, Cauchy, S. [†]Lie, and many other mathematicians. The fundamental existence theorem for the initial value problem, now called the Cauchy-Kovalevskaya theorem, was proved by S. Kovalevskaya in 1875 (→ 321 Partial Differential Equations (Initial Value Problems) B).

Because of their close connection with problems of physics, linear equations of the second order have been a chief object of research. Up to the 19th century, classification into [†]elliptic, [†]hyperbolic, and [†]parabolic types and the study of boundary and initial value problems for each of these types constituted the main part of the theory.

In the 20th century, more complicated

problems—[†]nonlinear problems appearing in the study of viscous or compressible fluids, or the study of equations of [†]mixed type in connection with supersonic flow—have emerged as important topics; and the newly developed techniques of functional analysis have brought about remarkable changes. Especially in the study of the [†]Schrödinger equations of quantum mechanics and of more general [†]evolution equations, this method has proved to be a powerful tool.

Finally, we should not fail to mention that the development of electronic computers has made it possible to obtain numerical solutions and to discover many important facts. [†]Numerical analysis is now becoming an indispensable part of the theory (→ 304 Numerical Solution of Partial Differential Equations).

References

- [1] E. L. Ince, Ordinary differential equations, Dover, 1956.
- [2] L. S. Pontryagin, Ordinary differential equations, Addison-Wesley, 1962. (Original in Russian, 1961.)
- [3] E. A. Coddington and N. Levinson, Theory of ordinary differential equations, McGraw-Hill, 1955.
- [4] P. Hartman, Ordinary differential equations, Wiley, 1964.
- [5] E. Hille, Lectures on ordinary differential equations, Addison-Wesley, 1969.
- [6] L. Bieberbach, Theorie der gewöhnlichen Differentialgleichungen auf funktionen theoretischen Grundlage dargestellt, Springer, 1953.
- [7] E. Hille, Ordinary differential equations in the complex domain, Wiley, 1976.
- [8] W. Wasow, Asymptotic expansions for ordinary differential equations, Interscience, 1965.
- [9] I. Kaplansky, An introduction to differential algebra, Hermann, 1957.
- [10] M. A. Naïmark, Linear differential operators, I, II, Ungar, 1967, 1968. (Original in Russian, 1954.)
- [11] R. Courant and D. Hilbert, Methods of mathematical physics I, II, Interscience, 1953, 1962.
- [12] P. R. Garabedian, Partial differential equations, Wiley, 1964.
- [13] I. G. Petrovskii, Lectures on partial differential equations, Interscience, 1954. (Original in Russian, 1950.)
- [14] L. Hörmander, Linear partial differential operators, Springer, 1963.
- [15] K. Yosida, Lectures on differential and integral equations, Wiley, 1960. (Original in Japanese, 1950.)

108 (XIX.10) Differential Games

A. Introduction

The study of differential games arose from the study of pursuit and evasion problems and various tactical problems. The first work was done by R. P. Isaacs [1] in a series of RAND Corporation memoranda that appeared in 1954. He applied to many illustrative examples the method of Hamilton-Jacobi differential equations (\rightarrow eq. (4) below). The heuristic results of Isaacs were made rigorous by W. H. Fleming [2], L. D. Berkovitz [3], A. Friedman [4, 5], and others.

B. Zero-Sum Two-Person Games

Suppose that there are two antagonists, each exerting partial control over the state of a system. One wishes to maximize a given payoff that is a functional of the state and the control exerted, while the other wishes to minimize this payoff. Let the state of a differential game at time t be represented by an n -dimensional vector $x(t) \in \mathbf{R}^n$. In a zero-sum differential game between two players I and II, we are given a system of n differential equations

$$dx/dt = f(t, x, u, v) \quad (1)$$

with an initial condition $x(\tau) = \xi \in \mathbf{R}^n$, where $u \in \mathbf{R}^p$ is chosen at each instant of time by player I and $v \in \mathbf{R}^q$ is chosen by player II. We assume that the function $f(t, x, u, v)$ is continuous in t and continuously differentiable on the entire (x, u, v) -space.

It is usually assumed that both players know the present state of the game and that they know how the game proceeds; that is, they know the system (1). Each player can take the state of the game into account in making his choice. Thus player I can let his choice of u be governed by a vector function $u(t, x)$ defined on D , where $D \subset [0, \infty) \times \mathbf{R}^n$ is a fixed region of the (t, x) -space. Similarly, player II can let his choice of v be governed by a vector function $v(t, x)$ defined on D .

A finite collection of subregions D_1, \dots, D_r of a region D is said to constitute a decomposition of D whenever the following conditions hold: (i) Each D_i ($i = 1, \dots, r$) is connected and has a piecewise smooth boundary; (ii) $D_i \cap D_j = \emptyset$ if $i \neq j$. A function defined on \bar{D} is said to be piecewise C^1 in x on \bar{D} if there is a decomposition of D such that on each D_i the function and all its derivatives with respect to x are continuous in (t, x) on \bar{D}_i . Let U and V be fixed closed subsets of \mathbf{R}^p and \mathbf{R}^q , respectively.

Let S_u denote the class of functions $u(t, x)$ that are piecewise C^1 in x on \bar{D} and have their range in U . Similarly, let S_v denote the class of functions $v(t, x)$ that are piecewise C^1 in x on \bar{D} and have their range in V .

Let $u \in S_u$ and $v \in S_v$, and consider the differential equation

$$dx/dt = f(t, x, u(t, x), v(t, x)), \quad (2)$$

subject to the initial condition

$$x(\tau) = \xi. \quad (3)$$

We say that a pair $(u, v) \in S_u \times S_v$ is **playable** if for every (τ, ξ) in D every solution of (2) satisfying (3) stays in D and reaches a terminal manifold F in finite time, where F is a smooth manifold contained in \bar{D} . Let $\Omega_u \subset S_u$, $\Omega_v \subset S_v$ be the maximal pair of subclasses such that each pair $(u, v) \in \Omega_u \times \Omega_v$ is playable. We call the functions in Ω_u and Ω_v the **strategies** for the players.

For each strategy pair (u, v) we can define a functional

$$J(\tau, \xi, u, v) = g(t_1, x(t_1)) + \int_{t_0}^{t_1} h(t, x(t), u(t, x(t)), v(t, x(t))) dt,$$

where $x(t)$ is a solution of (2) and (3) and t_1 is the first time that $(t, x(t))$ reaches the terminal manifold F . The functional J is called the **payoff**.

Let $(u^*, v^*) \in \Omega_u \times \Omega_v$ be a strategy pair. Suppose that for any $u \in \Omega_u$ and $v \in \Omega_v$, the inequalities

$$J(\tau, \xi, u, v^*) \leq J(\tau, \xi, u^*, v^*) \leq J(\tau, \xi, u^*, v)$$

hold for all $(\tau, \xi) \in D$. We say that (u^*, v^*) is a **saddle point** relative to the classes Ω_u and Ω_v . The function

$$W(t, x) = J(t, x, u^*, v^*)$$

defined on D is called the **value function** of the game. Berkovitz [3] proved that the value function $W(t, x)$ is continuous on D and continuously differentiable on each D_i and satisfies

$$\begin{aligned} & \max_{u \in U} [h(t, x, u, v^*) + W_x(t, x)f(t, x, u, v^*)] \\ &= \min_{v \in V} [h(t, x, u^*, v) + W_x(t, x)f(t, x, u^*, v)] \\ &= h(t, x, u^*, v^*) + W_x(t, x)f(t, x, u^*, v^*) \\ &= -W_t(t, x). \end{aligned} \quad (4)$$

Equation (4) is called the **Hamilton-Jacobi equation**.

Let $x^*(t; \tau, \xi)$ be the optimal trajectory corresponding to the saddle-point strategies (u^*, v^*) and resulting from an initial point $(\tau, \xi) \in D$. Then there exists an n -dimensional continuous

vector function $\lambda(t; \tau, \xi)$ such that the following hold [3]:

(i) The functions x^* and λ satisfy the system of differential equations

$$dx/dt = f(t, x, u^*(t, x), v^*(t, x)),$$

$$d\lambda/dt = -H_x(t, x, \lambda, u^*(t, x), v^*(t, x)),$$

where

$$H(t, x, \lambda, u, v) = h(t, x, u, v) + \lambda f(t, x, u, v).$$

(ii) If $x = x^*(t; \tau, \xi)$, $t \geq \tau$, then

$$W_x(t, x) = \lambda(t; \tau, \xi).$$

(iii) At $t = t_1$, the transversality condition

$$H \frac{\partial T}{\partial \sigma} + \frac{\partial g}{\partial t} \frac{\partial T}{\partial \sigma} + \frac{\partial g}{\partial x} \frac{\partial X}{\partial \sigma} - \lambda \frac{\partial X}{\partial \sigma} = 0$$

holds, where the terminal manifold F is given parametrically by the relations

$$t = T(\sigma), \quad x = X(\sigma),$$

σ ranging over a cube in some finite-dimensional Euclidean space.

(iv) For all $\tau \leq t \leq t_1$,

$$\max_{u \in U} \min_{v \in V} H(t, x^*(t), \lambda(t), u, v)$$

$$= \min_{v \in V} \max_{u \in U} H(t, x^*(t), \lambda(t), u, v)$$

$$= H(t, x^*(t), \lambda(t), u^*(t, x^*(t)), v^*(t, x^*(t))).$$

P. Varaiya and J. Lin [6] and Friedman [4, 5] have defined certain special classes of differential games, and have shown that under their definitions the games have nonzero value functions.

C. N-Person Differential Games

In a differential game between many players, the state vector $x(t) \in \mathbf{R}^n$ is governed by

$$dx/dt = f(t, x, u_1, \dots, u_N), \quad x(\tau) = \xi,$$

where $u_i \in \mathbf{R}^{p_i}$ is chosen at each instant of time by player i . Each u_i is constrained to lie in a fixed closed subset U_i of \mathbf{R}^{p_i} . Let S_i be the class of functions $u_i(t, x)$ that are piecewise C^1 in x on \bar{D} and have their range in U_i .

We say that an element $u = (u_1, \dots, u_N)$ of $S_1 \times \dots \times S_N$ is **playable** if, for every $(\tau, \xi) \in D$, every solution of the differential equation

$$dx/dt = f(t, x, u_1(t, x), \dots, u_N(t, x)), \quad x(\tau) = \xi, \quad (5)$$

stays in D and reaches a terminal manifold F in finite time. Let $\Omega_i \subset S_i$ ($i = 1, \dots, N$) be the maximal subclasses such that each element $u = (u_1, \dots, u_N) \in \Omega_1 \times \dots \times \Omega_N$ is playable. We call the functions $u_i \in \Omega_i$ ($i = 1, \dots, N$) the **strategies**. For each strategy N -tuple we define

a **payoff**

$$J_i(\tau, \xi, u_1, \dots, u_N) = g_i(t_1, x(t_1))$$

$$+ \int_{t_0}^{t_1} h_i(t, x(t), u_1(t, x(t)), \dots, u_N(t, x(t))) dt,$$

where $x(t)$ is a solution of (5) and t_1 is the first time that $(t, x(t))$ reaches the terminal manifold F . Each player i is to choose his strategy $u_i \in \Omega_i$ so as to maximize his own payoff J_i .

There are many definitions of "solution" for games involving more than two players. A strategy N -tuple $u^* = (u_1^*, \dots, u_N^*)$ is called an **equilibrium point** for the game if the inequalities

$$J_i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots)$$

$$\leq J_i(u_1^*, \dots, u_{i-1}^*, u_i^*, u_{i+1}^*, \dots) \quad (i = 1, \dots, N)$$

hold for any $u_i \in \Omega_i$, \dots , $u_N \in \Omega_N$. J. H. Case [9] has shown that the conclusions drawn for zero-sum two-person games also hold for N -person differential games.

References

- [1] R. P. Isaacs, Differential games, Wiley, 1965.
- [2] W. H. Fleming, The convergence problem for differential games II, Advances in game theory, Ann. Math. Studies 52, Princeton Univ. Press, 1964, 195–210.
- [3] L. D. Berkovitz, Necessary conditions for optimal strategies in a class of differential games and control problems, SIAM J. Control, 5 (1967), 1–24.
- [4] A. Friedman, On the definition of differential games and the existence of value and of saddle points, J. Differential Equations, 7 (1970), 69–91.
- [5] A. Friedman, Differential games, Wiley, 1971.
- [6] P. Varaiya and J. Lin, Existence of saddle points in differential games, SIAM J. Control, 7 (1969), 142–157.
- [7] H. W. Kuhn and G. P. Szegö (eds.), Differential games and related topics, North-Holland, 1971.
- [8] L. S. Pontryagin, On the theory of differential games, Russian Math. Surveys, 21 (1966), 193–246.
- [9] J. H. Case, Toward a theory of many player differential games, SIAM J. Control, 7 (1969), 179–197.

109 (VII.1) Differential Geometry

In differential geometry in the classical sense, we use differential calculus to study the prop-

erties of figures such as curves and surfaces in Euclidean planes or spaces. Owing to his studies of how to draw tangents to smooth plane curves, P. †Fermat is regarded as a pioneer in this field. Since his time, differential geometry of plane curves, dealing with curvature, †circles of curvature, †evolutes, †envelopes, etc., has been developed as a part of calculus. Also, the field has been expanded to analogous studies of space curves and surfaces, especially of †asymptotic curves, †lines of curvature, †curvatures and †geodesics on surfaces, and †ruled surfaces. C. F. †Gauss founded the theory of surfaces by introducing concepts of the †geometry on surfaces (*Disquisitiones circa superficiem curvas*, 1827). Gauss recognized the importance of the intrinsic geometry of surfaces, and it is generally agreed that differential geometry as it is known today was initiated by him. Thus differential geometry came to occupy a firm position as a branch of mathematics. The influence that differential-geometric investigations of curves and surfaces have exerted upon branches of mathematics, physics, and engineering has been profound. For example, E. Beltrami discovered an intimate relation between the geometry on a †pseudo-sphere and †non-Euclidean geometry. The study of †geodesics is a fertile topic deeply related to dynamics, the calculus of variations, and topology, on which there is excellent work by J. Hadamard, H. †Poincaré, P. Funk, G. D. Birkhoff, M. Morse, R. Bott, W. Klingenberg, and M. Berger, among others. The theory of minimal surfaces initiated by J. L. Lagrange was an application of the calculus of variations. At the early stages of development, G. Monge, J. B. M. C. Meusnier, A. M. Legendre, O. Bonnet, B. Riemann, K. Weierstrass, H. A. Schwarz, Beltrami, and S. Lie contributed to the theory. Weierstrass and Schwarz established its relationship with the theory of functions. J. A. Plateau showed experimentally that †minimal surfaces can be realized as soap films by dipping wire in the form of a closed space curve into a soap solution (1873). The Plateau problem, i.e., the problem of proving mathematically the existence of a minimal surface with prescribed boundary curve, was solved by Tibor Radó in 1930 and independently by J. Douglas in 1931. Although the relationship to function theory is lost for higher-dimensional minimal submanifolds, their study is intimately related to the calculus of variations and topology.

Euclidean geometry is a geometry belonging to F. Klein's Erlangen program (→ 137 Erlangen Program). For other geometries in the sense of F. Klein we may also consider the corresponding differential geometries. For instance, in †projective differential geometry we study

by means of differential calculus the properties of curves and surfaces that are invariant under projective transformations. This subject was studied by E. J. Wilczynski, G. Fubini, and others; †affine differential geometry and †conformal differential geometry were studied by W. Blaschke and others (→ 110 Differential Geometry in Specific Spaces).

Influenced by Gauss's geometry on surfaces, in his inaugural address at Göttingen in 1854 Riemann advocated an intrinsic differential geometry completely independent of embeddings (*Über die Hypothesen, welche der Geometrie zugrund liegen*; *Werke*, 2nd ed. 1892, 272–287) (→ 364 Riemannian Manifolds). Removing the restriction to two dimensions and considering abstract manifolds of dimension n , he introduced what is now known as the Riemannian metric; actually he considered the more general metrics that had formed the subject matter of the dissertation of P. Finsler in 1918 (→ 152 Finsler Spaces). Riemannian geometry includes Euclidean and non-Euclidean geometry as special cases, and is important for the great influence it exerted on geometric ideas of the 20th century. Under the influence of the algebraic theory of invariants, Riemannian geometry was then studied as a theory of invariants of quadratic †covariant tensors by E. B. Christoffel, C. G. Ricci, and others. Riemannian geometry attracted wide attention after A. Einstein applied it to the †general theory of relativity in 1916.

In the same year, T. Levi-Civita introduced the notion of †Levi-Civita parallelism, which contributed greatly to the clarification of geometric properties of Riemannian spaces. Observing parallelism to be an affine-geometric concept, H. Weyl and A. S. Eddington developed a theory of Riemannian spaces “affinely” based on the notion of parallelism without using metrical methods. Such a geometry is called a geometry of an affine connection (→ 80 Connections).

Every straight line in a Euclidean space has the property that all tangents to the line are parallel. In a space with an affine connection, we may define a family of curves called †paths as an analog of straight lines. Such curves are solutions of a system of ordinary differential equations of the second order of a certain type. Coefficients of such differential equations determine a parallelism and hence an affine connection. H. Weyl discovered transformations of coefficients that leave the family of paths invariant as a whole, namely, projective transformations of an affine connection. A geometry that aims to study properties of paths or affine connections that are invariant under these transformations is called a **projective geometry of paths**. Such geometry was

studied by L. P. Eisenhart, O. Veblen, and others. The concept of projective connections was an outcome of such studies. Similarly, the concept of conformal connections was developed from the consideration of [†]conformal transformations of Riemannian spaces.

These geometries cannot in general be regarded as geometries in the sense of Klein. Actually, any one of these geometries generally has no transformations that correspond to [†]congruent transformations of geometries in the sense of Klein; even if it has such transformations, they do not act transitively on the space. Thus geometries are naturally divided into two categories, one consisting of geometries in the sense of Klein (based on the group concept) and the other of geometries based on Riemann's idea. Under such circumstances, E. Cartan unified the thoughts of Klein and Riemann from a higher standpoint and constructed his theory of connections in a series of papers published between 1923 and 1925. He developed the theory of affine, projective, and conformal connections from a viewpoint consistent with that of Klein. Just as each tangent space of a Riemannian manifold is viewed as a Euclidean space, an affine connection regards the tangent space at each point as an affine space and develops it onto the tangent space at an infinitesimally nearby point. In discussing projective connections Cartan attached a projective space to each point of a manifold as an infinitesimal approximation, and similarly for conformal connections. More generally, he attached to each point of a manifold a fixed Klein space, i.e., a homogeneous space of a Lie group, called the structure group. Thus Cartan introduced the concept of fiber bundle (\rightarrow 147 Fiber Bundles). Then he defined a connection as a development of the fiber, i.e., the generalized tangent space, at each point onto the fiber at an infinitesimally nearby point (\rightarrow 80 Connections B). If G is the group of congruent transformations in Euclidean space, a manifold with connection having G as its structural group is called a **manifold with Euclidean connection**. Among manifolds with Euclidean connection, Riemannian manifolds are characterized as those without [†]torsion. If we take the group of congruent transformations of projective (conformal) geometry as G , we have manifolds with projective (conformal) connection in the sense of Cartan. Among these, there are remarkable ones called manifolds with normal projective (or conformal) connections, which are essentially the same as the ones studied by Veblen and others. Cartan's idea had a profound influence on modern differential geometry. The method of moving frames, created by G. Darboux and extensively used by Cartan in his theory of connections, was a

forerunner of the theory of fiber bundles. Combining [†]Grassmann algebra with differential calculus, Cartan developed a powerful computational tool known as calculus of differential forms (\rightarrow 105 Differentiable Manifolds Q). Differential forms have become indispensable in topology, algebraic geometry, and in studies of functions of several complex variables, as well as in differential geometry.

The work of Lie on transformation groups also had a profound influence on Cartan. The latter's work on [†]Lie groups, particularly on simple Lie groups, and on differential geometry culminated in 1926 in his discovery of [†]Riemannian symmetric spaces. These spaces are natural generalizations of the spherical surface and the unit disk in the complex plane with Poincaré metric, and play essential roles in unitary representation theory and in other areas of mathematics.

A tangent line to a curve C at a point P of C is the limit line of the line PQ , where Q is a point on C approaching P ; hence we can define it locally. A concept (or property), such as this, that can be defined in an arbitrary small neighborhood of a point of a given figure or a space is called a **local concept** (or **local property**) or a concept (or property) **in the small**. In the early stages of the development of differential geometry, differential calculus was the main tool of study, so most of the results were local. On the other hand, a concept (or property) that is defined in connection with a whole figure or a whole space is called **global** or **in the large**. In modern differential geometry, the study of relations between local and global properties has attracted the interest of mathematicians. This view was emphasized by Blaschke, who worked on the differential geometry of [†]ovals and [†]ovaloids. The study of [†]rigidity of ovaloids by S. Cohn-Vossen belongs in this category, and many works on geodesics and minimal surfaces were done from this standpoint.

From the viewpoint of modern mathematics, the basic concepts on which we construct Riemannian geometry and geometries of connections are global concepts of [†]differentiable manifolds. However, in Riemann's time the theory of [†]Lie groups and topology were not yet developed; consequently, Riemannian geometry remained a local theory. In 1925, H. Hopf began to study the relations between local differential-geometric structures and the topological structures of Riemannian spaces. However, except for the work of Cartan, Hopf, and a few others, differential geometry in the 1920s was still largely concerned with surfaces in the 3-dimensional Euclidean space or local properties of Riemannian manifolds, and with affine, projective, and conformal connections.

Gradually the concept of differentiable manifolds was clarified, the global theory of Lie groups made progress, and topology developed; and the trend toward global differential geometry began slowly in the early 1930s. The dissertation of G.-W. de Rham published in 1931 showed that the cohomology of a manifold can be computed in terms of differential forms (\rightarrow 105 Differentiable Manifolds R). His theorem provides the theoretical foundation for expressing cohomological invariants of a manifold in terms of differential geometric invariants. In a series of papers immediately following de Rham's, W. V. D. Hodge established that, on a compact Riemannian manifold, every r -dimensional cohomology class can be uniquely represented by a harmonic form of degree r (\rightarrow 194 Harmonic Integrals).

An important class of complex manifolds with compatible Riemannian metric was discovered by J. A. Schouten, D. van Dantzig, and E. Kähler around 1929–1932. This class of manifolds, called Kähler manifolds today, comprises the projective algebraic manifolds. Hodge's theory of harmonic integrals is most effective when applied to compact Kähler manifolds, (\rightarrow 232 Kähler Manifolds).

The most celebrated global theorem in classical differential geometry of surfaces is the Gauss-Bonnet formula (1848) (\rightarrow 364 Riemannian Manifolds D). The formula was generalized to closed hypersurfaces of Euclidean space by Hopf in 1925, to closed submanifolds of Euclidean space by C. B. Allendoerfer and W. Fenchel in 1940, and finally to arbitrary closed Riemannian manifolds by Allendoerfer and A. Weil in 1943. But the simple proof given by S. S. Chern in 1944 contained the notion of transgression, which has become essential in the theory of characteristic classes (\rightarrow 56 Characteristic Classes). The discovery of Pontryagin classes for Riemannian manifolds (1944) and Chern classes for Hermitian manifolds (1946) culminated in the index theorem and the Riemann-Roch theorem of F. E. P. Hirzebruch, and finally in the Atiyah-Singer index theorem.

A simple but fruitful idea of S. Bochner, relating harmonic forms to curvature, established vanishing theorems for harmonic forms of Riemannian manifolds and for holomorphic forms of Kähler manifolds under suitable positivity conditions for curvature. His idea has led to the vanishing theorems of K. Kodaira and others (\rightarrow 232 Kähler Manifolds D).

The work of C. Ehresmann in 1950 on connections in principal fiber bundles established a solid foundation to Cartan's theory of connections. Gauge theory in physics is largely

based on the theory of connections in principal fiber bundle.

R. H. Nevanlinna's value distribution theory and its subsequent generalization by Chern and others can be best described in differential-geometric terms.

Differentiable manifolds are currently objects of research in both differential geometry and differential topology. While topology studies manifolds per se, differential geometry may be considered as the study of differentiable manifolds equipped with geometric structures, such as metric tensors, connections, (almost) complex structures, and various other tensors. Through these geometric structures, differential geometry enjoys close contact with many branches of mathematics. From its early days, differential geometry has had close ties to topology (as exemplified by the Gauss-Bonnet formula) and to partial differential equations and analytic functions (through, e.g., the study of minimal surfaces). The bonds with topology were strengthened by Morse theory and, more recently, by the theory of characteristic classes. In the most recent proof of the Atiyah-Singer index theorem, differential geometry is an important intermediary between topology and analysis. Differential geometry and algebraic geometry have enriched each other through Kähler manifolds. The theory of functions of several complex variables also has points of contact with differential geometry, such as value distribution theory and Cauchy-Riemann structures. Contact and symplectic structures are basic to mechanics. Lorentz manifolds and connections in principal bundles are essential mathematical tools in the general theory of relativity and in gauge theory. Topics such as minimal submanifolds, manifolds of positive curvature, and closed geodesics are active and important areas of research belonging to Riemannian geometry proper; at the same time, differential geometry provides a language and methods that are important in wider areas of mathematics.

References

- [1] L. P. Eisenhart, *A treatise on the differential geometry of curves and surfaces*, Ginn, 1909.
- [2] G. Darboux, *Leçons sur la théorie générale des surfaces*, Gauthier-Villars, I, second edition, 1914; II, second edition, 1915; III, 1894; IV, 1896.
- [3] W. Blaschke, *Vorlesungen über Differentialgeometrie*, Springer, I, third edition, 1930; II, 1923; III, 1929 (Chelsea, 1967).
- [4] L. P. Eisenhart, *Riemannian geometry*, Princeton Univ. Press, 1926.

- [5] E. Cartan, *Leçons sur la géométrie des espaces de Riemann*, Gauthier-Villars, second edition, 1946; English translation, *Geometry of Riemannian spaces*, Math. Sci. Press, 1983.
- [6] J. A. Schouten, *Ricci-calculus*, Springer, second edition, 1954.
- [7] O. Veblen and J. H. C. Whitehead, *The foundations of differential geometry*, Cambridge Tracts, 1932.
- [8] S. S. Chern, *Topics in differential geometry*, Lecture notes, Inst. for Adv. Study, Princeton, 1951.
- [9] W. V. D. Hodge, *The theory and applications of harmonic integrals*, Cambridge Univ. Press, second edition, 1952.
- [10] S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, Interscience, I, 1953; II, 1969.
- [11] E. Cartan, *Oeuvres complètes I–III*, Gauthier-Villars, 1952–1955; second edition, CNRS, 1984.
- [12] A. L. Besse, *Manifolds all of whose geodesics are closed*, Springer, 1978.
- [13] M. Berger, P. Gauduchon, and E. Mazet, *Le spectre d'une variété Riemannienne*, Lecture notes in math. 194, Springer, 1971.
- [14] D. Gromoll, W. Klingenberg, and W. Meyer, *Riemannsche Geometrie im Großen*, Lecture notes in math. 55, Springer, 1968.
- [15] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Academic Press, 1978.
- [16] S. Kobayashi, *Transformation groups in differential geometry*, Springer, 1972.
- [17] T. Levi-Civita, *The absolute differential calculus*, Blackie, 1926.
- [18] J. W. Milnor, *Lectures on Morse theory*, Ann. Math. Studies 51, Princeton Univ. Press, 1963.
- [19] R. Osserman, *A survey of minimal surfaces*, Van Nostrand, 1969.
- [20] G. de Rham, *Variétés différentiables*, Hermann, 1955.
- [21] A. Weil, *Introduction à l'étude des variétés kähleriennes*, Hermann, 1958.
- [22] K. Yano and S. Bochner, *Curvature and Betti numbers*, Ann. Math. Studies 32, Princeton Univ. Press, 1953.
- [23] Y. Matsushima, *Differentiable manifolds*, Dekker, 1972.

110 (VII.17) Differential Geometry in Specific Spaces

A. The Method of Moving Frames

The main theme of this article is the theory of surfaces (i.e., submanifolds) in a differentiable

manifold V on which a t Lie transformation group G acts.

If a t Lie group G of dimension r acts t transitively on a space G_* and the t stability group for any point of G_* consists of the identity element only, then G_* is called the **group manifold** of G , and an element of G_* is called a **frame**. If f_0 is a fixed frame, then the mapping $a \rightarrow af_0 \in G_*$ ($a \in G$) gives a t diffeomorphism of G to G_* . Let $I(G)$ be the set of all t differential forms ω of degree 1 on G_* that are invariant under transformations of G . Then $I(G)$ is a linear space of dimension r and is the t dual space of the Lie algebra \mathfrak{g} of G . A basis $\{\omega_\lambda | 1 \leq \lambda \leq r\}$ of $I(G)$ is called a set of **relative components** of G . The **structure equations** hold:

$$d\omega_\lambda = \frac{1}{2} \sum_{\mu, \nu=1}^r c_{\lambda\mu\nu} \omega_\mu \wedge \omega_\nu,$$

where $c_{\lambda\mu\nu}$ are t structure constants of the Lie algebra \mathfrak{g} .

Let G be a Lie transformation group of a space E . Then a G -invariant submanifold of E on which G acts transitively is called an **orbit**. Each point $y \in E$ determines an orbit containing y . When there exist parameters k_j ($1 \leq j \leq t$) such that any G -invariant on E is a function of k_1, \dots, k_t , then these parameters are called the **fundamental invariants** of E . Let H ($\subset G$) be the stability group at a point y_0 on an orbit M ; then M is identified with the t homogeneous space G/H by the diffeomorphism $\varphi: M \rightarrow G/H$, $\varphi(ay_0) = aH$ ($a \in G$). Furthermore, a t principal fiber bundle (G_*, M, H, τ) is determined by the projection $\tau: G_* \rightarrow M$, $\tau(af_0) = ay_0$ ($a \in G$). The t fiber H_y on a point $y \in M$ is a group manifold of H . H_y is called the **family of frames** on y , and an element of H_y is a frame on y . Local coordinates θ_μ ($1 \leq \mu \leq s$) of the group H are called the **secondary parameters** and are used to indicate frames in H_y . When H is not connected, let H^0 be the connected component of the identity of H and \tilde{M} be the t covering manifold G/H^0 of M . An element $\tilde{y} \in \tilde{M}$ over $y \in M$ is called an **oriented element**. Now assume that the group H is connected. Then the family of frames H_y on each $y \in M$ is given as an t integral manifold of a t completely integrable system of total differential equations $\pi_i = 0$ ($1 \leq i \leq r-s$, $\pi_i \in I(G)$) on the group manifold G_* . Here the π_i are linearly independent and are called the **horizontal components** of M . The π_i are linear combinations of the relative components ω_λ of G , and their coefficients are generally functions of the fundamental invariants k_j . For simplicity, we assume that the relative components $\{\omega_\lambda\}$ are chosen such that the horizontal components π_i and components ω_α ($r-s < \alpha \leq r$) are linearly independent. Then the ω_β ($1 \leq \beta \leq s$) are called the **secondary com-**

ponents. The differentials $d\theta_\mu$ of the secondary parameters are linear combinations of the ω_μ . Furthermore, let $\{x_\sigma\}$ ($1 \leq \sigma \leq n$) be local coordinates of E ; then the differentials dx_σ are linear combinations of the differentials dk_j and the horizontal components π_i .

Let G be a Lie transformation group of a space V . We regard two m -dimensional surfaces W_1 and W_2 passing through $x \in V$ as equivalent if they have a \dagger contact of order p at x . Then an equivalence class of submanifolds is called a **contact element** of order p at the point x . Let E_p be the set of all contact elements of order p at x , where x runs over all the points in V . A contact element of order p naturally determines a contact element of order $p-1$, and we denote this correspondence by $\psi: E_p \rightarrow E_{p-1}$. Thus we obtain the series of correspondences

$$V = E_0 \xleftarrow{\psi} E_1 \leftarrow \dots \leftarrow E_{p-1} \xleftarrow{\psi} E_p \leftarrow \dots,$$

where a contact element of order 0 is identified with a point of V . Since a transformation on the space V induces a transformation on E_p , G is also a Lie transformation group of E_p , and this transformation commutes with the mapping ψ . The fundamental invariants k_j of E_p are said to be of order p . We use similar terminology (such as frames of order p , etc.) throughout this article. The fundamental invariants k_j of order p ($1 \leq j \leq t_p$) can be chosen such that they contain the fundamental invariants k_i ($1 \leq i \leq t_{p-1}$) of order $p-1$. The additional $t_p - t_{p-1}$ invariants k_α ($t_{p-1} < \alpha \leq t_p$) are called the **invariants of order p** . The family H_y^p ($y \in E_p$) of frames of order p can be chosen such that H_y^p is contained in the family H_z^{p-1} ($z = \psi y \in E_{p-1}$) of frames of order $p-1$. If necessary, the family H_y^p of frames of order p can be made connected by defining an **orientation** of contact elements of order p . Furthermore, the horizontal components π_j ($1 \leq j \leq r - s_p$) of order p can be chosen such that they contain the horizontal components π_i ($1 \leq i \leq r - s_{p-1}$). The additional $s_{p-1} - s_p$ components π_α ($r - s_{p-1} < \alpha \leq r - s_p$) are called the **principal components of order p** .

Let W be an m -dimensional surface of a space V . The contact element of order p (≥ 0) is determined at every point of W and expressed by the family of frames of order p and the values of invariants of orders less than or equal to p . Let $\{u_i\}$ ($1 \leq i \leq m$) be local coordinates on W . Then the differentials du_i are given as linear combinations of linearly independent differential forms π_i ($1 \leq i \leq m$), where the π_i , called the **basic components** of W , are certain linear combinations of the differentials of the fundamental invariants of V and of the horizontal components of orbits of V . Let $F^p(W)$ be the set of all families of frames of order p ; then

$F^p(W)$ depends on m parameters u_i and s_p secondary parameters θ_μ of order p . On the space $F^p(W)$, the differentials of invariants of orders less than or equal to $p-1$ and the principal components of orders less than or equal to $p-1$ are linear combinations of the basic components, whose coefficients are functions of the invariants of orders less than or equal to p . The differentials of invariants of order p and the principal components of order p are linear combinations of the basic components:

$$dk_\alpha = h_{\alpha 1} \pi_1 + \dots + h_{\alpha m} \pi_m, \quad t_{p-1} < \alpha \leq t_p,$$

$$\pi_\alpha = b_{\alpha 1} \pi_1 + \dots + b_{\alpha m} \pi_m, \quad r - s_{p-1} < \alpha \leq r - s_p,$$

where the coefficients $h_{\alpha i}$, $b_{\alpha i}$ are functions of the invariants of orders less than or equal to p and, in general, the secondary parameters θ_μ of order p . These coefficients are called the **coefficients of order p** . Let Γ_p be a subgroup of G preserving a family of frames of order p and D_p be a space whose coordinates are coefficients $(h_{\alpha i}, b_{\alpha i})$ of order p . Then Γ_p acts on D_p as a transformation group. Knowledge of the properties of contact elements of order less than or equal to p can be utilized to obtain information about the invariants of order $p+1$, etc. In fact, if we can choose in the Γ_p -space D_p a subspace C_p that intersects each orbit in D_p at one and only one point, then in general the secondary parameters of order p associated with the points in C_p correspond to the frames of order p , and the parameters associated with the points in C_p are the invariants of order $p+1$. The restrictions of the coefficients of order p to C_p are functions of the invariants of orders less than or equal to $p+1$; they are independent of the secondary parameters of order p . Thus the frames of order $p+1$ and the invariants of orders less than or equal to $p+1$ determine the contact elements of order p of W and their differentials; generally, the latter can be utilized to determine the contact elements of order $p+1$.

This process of obtaining information of "order $p+1$ " utilizing a suitable subspace C_p of D_p is the so-called general **method of moving frames**. However, the surface W may contain points for which the general method does not apply. Actually, there are surfaces W for which the method does not apply for any point in W . Thus various methods of moving frames are necessary to cope with different kinds of surfaces. In the actual application of the method of moving frames, we use certain devices that help to simplify the calculations. In fact, an infinitesimal transformation $(\delta h_{\alpha i}, \delta b_{\alpha i})$ of the group Γ_p acting on the space D_p is expressed as a linear combination of the secondary components of order p ; this expression is easily obtained by means of the structure equations of

G. The group Γ_{p+1} is a subgroup of Γ_p fixing every point of the subspace C_p , and its infinitesimal transformation is such that $\delta h_{xi}=0$, $\delta b_{xi}=0$. The secondary components of order $p+1$ are immediately obtained from the equations for dk_x and π_x . Furthermore, when $m \geq 2$, the condition for the principal components of every order to satisfy the structure equations of G is essential to the problem of the existence of (m -dimensional) surfaces.

As we apply the method of moving frames consecutively to a surface W , we eventually arrive at the order q having the following properties: The families of frames of order $q+1$ coincide with those of order q , and the invariants k_β of order $q+1$ are expressed as functions $\varphi_\beta(k_x)$ of invariants k_x of order less than or equal to q . In this case, the families of frames of orders $q+j$ ($j \geq 1$) are all equal, and the invariants of orders $q+j$ are partial ($j-1$)-derivatives of functions $\varphi_\beta(k_x)$. The family of frames of order q is called the **Frenet frame**. The **differential invariants** on a surface are defined to be differential forms generated by the basic components and the invariants of each order.

Specifically, assume that the group G is an analytic transformation group of V , and the m -dimensional surfaces W_1, W_2 are analytic. Then there exists an element g of G such that $gW_1 = W_2$ if and only if W_1 and W_2 are of the same kind and have the same relations among the invariants of orders less than or equal to $q+1$. These relations are called the **natural equations** of the surface. The theory of surfaces based on the analysis of the natural equations of surfaces is called **natural geometry**. The **reduction formula** can be obtained by utilizing the Frenet frame; it gives the equation of the surface in the form of power series containing the invariants of each order.

Various results are known concerning the theory of surfaces of the spaces V_1, V_2 sharing the same transformation group G . We also have a theory of **special surfaces** whose invariants satisfy specific functional relations. Furthermore, we have problems concerning the **deformation of a surface** (preserving some differential invariants). Actually, the theory of surfaces of dimension m other than curves and hypersurfaces is in general quite difficult. The methods of tensor calculus can be applied to the study of surfaces. The theory of \dagger connections can be considered to be an outgrowth of the study of surfaces by means of the method of moving frames and tensor calculus.

B. Projective Differential Geometry

The rudiments of differential geometry subordinated to the \dagger projective transformation

group, or **projective differential geometry**, can be found in the *Theory of surfaces* by J. G. Darboux. The subject has been systematically studied by H. G. H. Halphen, E. J. Wilczynski, and G. Fubini. The Fubini theory was enriched substantially by E. Cartan, E. Čech, E. Bompiani, and J. Kanitani.

In this section we consider a surface S in a 3-dimensional projective space. Let $A(u^1, u^2)$ (u^1, u^2 are parameters on S) be a point of S , and associate with A all the frames $[A, A_1, A_2, A_3]$ ($|A, A_1, A_2, A_3| = 1$), where A_1, A_2, A_3 are points of the tangent plane to S at A . A family of such frames is called the **family of frames of order 1**, and we express its differential by

$$dA_\alpha = \sum_{\beta=0}^3 \omega_\alpha^\beta A_\beta, \quad \alpha=0, 1, 2, 3, \quad A_0 = A.$$

The ω_α^β are \dagger Pfaffian forms that depend on two principal parameters determining the origin A and ten secondary parameters determining the frame. We have $\omega_0^0 + \omega_1^1 + \omega_2^2 + \omega_3^3 = 0$, $\omega_0^3 = 0$. Furthermore, $\omega^1 = \omega_0^1$, $\omega^2 = \omega_0^2$ are independent of each other and depend on the principal parameters only. Let z^1, z^2, z^3 be \dagger nonhomogeneous coordinates with respect to a frame of order 1. Then in a neighborhood of the origin, S is expressed by $z^3 = \sum_{r=2}^\infty f_r$, where the f_r are homogeneous functions of degree r with respect to z^1, z^2 .

If we write $f_2 = (a_0(z^1)^2 + 2a_1 z^1 z^2 + a_2(z^2)^2)/2$, then it follows from the structure equations of the projective transformation group that $\omega_1^3 = a_0 \omega^1 + a_1 \omega^2$, $\omega_2^3 = a_1 \omega^1 + a_2 \omega^2$. If we put $\varphi_2 = a_0(\omega^1)^2 + 2a_1 \omega^1 \omega^2 + a_2(\omega^2)^2$, then a curve on S defined by $\varphi_2 = 0$ is called the **asymptotic curve** and its tangent the **asymptotic tangent**. At any point of this curve, the plane tangent to S is in contact of order 2 with this curve, and there are in general two asymptotic curves through any point of S . Equations of the asymptotic tangent at A are given by $z^3 = 0$, $f_2 = 0$. A point of S at which the asymptotic tangents coincide is called a **parabolic point**. If every point of S is parabolic, then S is a \dagger developable surface, and the general theory is not applicable to such a surface.

Among the family of frames of order 1, a frame satisfying $a_0 = a_2 = 0$, $a_1 = 1$ is called the **frame of order 2**. For this frame, the straight lines $\overline{AA_1}, \overline{AA_2}$ are asymptotic tangents. With respect to this frame, if $f_3 = -(b_0(z^1)^3 + 3b_1(z^1)^2 z^2 + 3b_2 z^1(z^2)^2 + b_3(z^2)^3)/3$, then $\omega_1^2 = b_0 \omega^1 + b_1 \omega^2$, $-\omega_0^0 + \omega_1^1 + \omega_2^2 - \omega_3^3 = 2(b_1 \omega^1 + b_2 \omega^2)$, $\omega_2^1 = b_2 \omega^1 + b_3 \omega^2$, and the quadric surface $z^3 = z^1 z^2 - z^3(b_1 z^1 + b_2 z^2 + p z^3)$ (with p arbitrary) is called **Darboux's quadric** at A , an especially interesting one among contact quadrics of S . **Darboux's curve** is a curve on S such

that Darboux's quadric is in contact of order 3 at any point of it. Its tangent is called **Darboux's tangent** and is given by $z^3=0$, $b_0(z^1)^3 + b_3(z^2)^3=0$.

We have $b_0 b_3 \neq 0$, except in the case of ruled surfaces. We take special frames of order 2 determined by $b_1=b_2=0$, $b_0=b_3=1$ and call them the **frames of order 3**. If a frame of order 3 satisfies $f_4 = -(c_0(z^1)^4 + 4c_1(z^1)^3 z^2 + 6(c_2-1)(z^1 z^2)^2 + 4c_3 z^1 (z^2)^3 + c_4(z^2)^4)/12$, then $\omega_0^0 - 2\omega_1^1 + \omega_2^2 = c_0 \omega^1 + c_1 \omega^2$, $\omega_3^2 - \omega_1^0 = c_1 \omega^1 + c_2 \omega^2$, $\omega_3^1 - \omega_2^0 = c_2 \omega^1 + c_3 \omega^2$, $\omega_0^0 + \omega_1^1 - 2\omega_2^2 = c_3 \omega^1 + c_4 \omega^2$. With respect to this family of frames of order 3, $((\omega^1)^3 + (\omega^2)^3)/2\omega^1 \omega^2$ is an invariant associated with two neighboring points of S , called the **projective line element**. Also with respect to this frame, two straight lines $\overline{AA_3}$, $\overline{A_1 A_2}$ are polar with respect to Darboux's quadric.

Among the families of frames of order 3, a frame satisfying $c_1=c_2=c_3=0$ is called a **frame of order 4**. With respect to this frame, there exist λ, μ, ν, ρ such that $\omega_1^0 = \lambda \omega^1 + \mu \omega^2$, $\omega_2^0 = \nu \omega^1 + \rho \omega^2$, $\omega_3^0 = \rho \omega^1 + \lambda \omega^2$. Hence if we put $c_0 = -3a$, $c_4 = -3b$, it follows that

$$(\omega_a^b) = \begin{pmatrix} \tau_0^0 & \omega^1 & \omega^2 & 0 \\ \omega_1^0 & \tau_1^1 & \omega^1 & \omega^2 \\ \omega_2^0 & \omega^2 & -\tau_1^1 & \omega^1 \\ \omega_3^0 & \omega_2^0 & \omega_1^0 & -\tau_0^0 \end{pmatrix},$$

$$\tau_0^0 = -(3/2)(a\omega^1 + b\omega^2), \quad \tau_1^1 = (1/2)(a\omega^1 + b\omega^2).$$

Thus the frame of order 4 is the Frenet frame and is attached to every point of S . This frame is called the **normal frame**, and the invariants $a, b, \lambda, \mu, \nu, \rho$ are called the **fundamental differential invariants**. The straight lines $\overline{AA_3}$ and $\overline{A_1 A_2}$ associated with the normal frame are called **directrices of Wilczynski** of the first and second kind, respectively. With respect to the normal frame, S is expressed by

$$z^3 = z^1 z^2 - ((z^1)^3 + (z^2)^3)/3 + (a(z^1)^4 + b(z^2)^4)/4 + (z^1 z^2)^2/2 + \dots$$

A necessary and sufficient condition for two surfaces S, \bar{S} to be projectively equivalent is that there be normal frames having the same ω^1, ω^2 and the same six fundamental differential invariants. For six quantities $a, b, \lambda, \mu, \nu, \rho$ to be fundamental differential invariants of a surface, they must satisfy a certain condition of existence [6].

A frame of order 1 such that $\overline{AA_3}$ and $\overline{A_1 A_2}$ are polar with respect to Darboux's quadric is called **Darboux's frame**. With respect to this frame also, a theory of surfaces has been established.

Consider a pointwise correspondence between two surfaces S, \bar{S} , and denote by $\bar{A} \in \bar{S}$ the point corresponding to $A \in S$. If there exists a projective transformation φ that transforms

A into \bar{A} and the image $\varphi(S)$ is in contact of order 2 with \bar{S} at \bar{A} , then the pointwise correspondence is called a **projective deformation**. A necessary and sufficient condition for the existence of a projective deformation between two surfaces is that these surfaces have the same projective line element [6]. A ruled surface is projectively deformable only to a ruled surface. Given an arbitrary surface S , it is generally impossible to find a surface that is different from S and projectively deformable to S ; some conditions must be satisfied [6, 8].

Let $p^{01}, p^{02}, p^{03}, p^{12}, p^{13}, p^{23}$ be Plücker coordinate of a straight line in a 3-dimensional projective space P^3 . Then we have $p^{01} p^{23} - p^{02} p^{13} + p^{03} p^{12} = 0$, and there is a one-to-one correspondence between the ratios of $\{p^{ij}\}$ and straight lines (\rightarrow 90 Coordinates B). If the p^{ij} are regarded as homogeneous coordinates of a 5-dimensional projective space P^5 , then the previous equation defines a hyperquadric Q in P^5 . Thus there is a one-to-one correspondence between points of Q and straight lines in P^3 . A curve on Q corresponds to a set of one-parameter families of straight lines, or a ruled surface. Sets of 2-parameter or 3-parameter families of straight lines corresponding to surfaces of 2 or 3 dimensions on Q in P^5 are called **congruences of lines** or **complexes of lines**, respectively. Thus by using a theory of surfaces in P^5 , it is possible to establish the theory of congruences and complexes [2, 6, 8], which is an important part of projective differential geometry.

Specifically, if the surface is either a curve or a hypersurface, there are numerous interesting results [2, 4, 6].

C. Affine Differential Geometry

The theme of general affine differential geometry is the study of differential-geometric properties of a point or set of points in a space that are invariant under the action of the affine transformation group. **Affine differential geometry** is the study of the properties invariant under the action of the equivalent affine transformation group, i.e., a subgroup of the affine transformation group formed by elements sending (x_i) to (\bar{x}_i) such that

$$\bar{x}_i = a_i + \sum_{j=1}^n a_{ij} x_j, \quad i = 1, \dots, n, \quad \det(a_{ij}) = 1.$$

The latter transformation leaves invariant the volume surrounded by an oriented closed hypersurface. The method of moving frames is effective in affine differential geometry.

Let C be a plane curve, and associate with any point $A = A(t)$ of C a family of frames $[A, e_1, e_2]$, where the area of the parallelogram

Differential Geometry in Specific Spaces

determined by the two vectors $\mathbf{e}_1, \mathbf{e}_2$ is equal to 1. This frame is called the **frame of order 0**. Its differential is expressed by

$$dA = \sum_{s=1}^2 \omega^s \mathbf{e}_s, \quad d\mathbf{e}_r = \sum_{s=1}^2 \omega_r^s \mathbf{e}_s, \quad r=1, 2,$$

$$\omega_1^1 + \omega_2^2 = 0.$$

The frames of order 1, 2, and 3 are characterized by $\omega^2 = 0$; $\omega^2 = 0, \omega_1^1 = \omega^1$; and $\omega^2 = 0, \omega_1^2 = \omega^1, \omega_1^1 = 0$, respectively. Then a frame of order 3 can be associated with each point of C and coincides with the Frenet frame. We call $\omega^1 = d\sigma$ the **affine arc element**; the **affine curvature** κ is defined by $\omega_2^1 = -\kappa d\sigma$. Then the Frenet formula is given by

$$dA = d\sigma \mathbf{e}_1, \quad d\mathbf{e}_1 = d\sigma \mathbf{e}_2, \quad d\mathbf{e}_2 = -\kappa d\sigma \mathbf{e}_1.$$

With respect to this frame, C is expressed as

$$y = x^2/2 + \kappa x^4/8 + (d\kappa/d\sigma)x^5/40 + \dots$$

Further, $d\sigma$ and κ are given analytically by

$$d\sigma = |dA, d^2A|^{1/3}, \quad \kappa = |d^2A/d\sigma^2, d^3A/d\sigma^3|,$$

where $|M, N| = \det(M, N)$. M, N are column vectors with two entries. We call σ the **affine arc length**. The straight line on which \mathbf{e}_2 is situated is called the **affine normal**, the diameter of the parabola osculating C at A . If κ is constant, then C is a conic section. Furthermore, C is an ellipse, hyperbola, or parabola according as the constant κ is positive, negative, or zero. In affine geometry, parabolas play a role similar to that played by straight lines in Euclidean geometry.

There are numerous results concerning the theory of skew curves and surfaces [1]. Concerning the theory of skew curves, results on **affine length**, **affine curvature**, **affine torsion**, **affine principal normals**, and **affine binormals** are similar to those in Euclidean geometry. The affine transformation group is situated between the projective transformation group and the congruent transformation group and hence has properties analogous to theirs. The theory of surfaces has a character similar to that of projective differential geometry [1]. We may also consider the variation of the affine area of a surface surrounded by a closed skew curve C . We call the extremal surface the **affine minimal surface**. W. Blaschke and others obtained many results on the global properties of such surfaces.

D. Conformal Differential Geometry

Let S^n be a conformal space of dimension n , and associate with each point $A_0 \in S^n$ a frame $\mathfrak{R}[A_0, A_1, \dots, A_n, A_\infty]$ of the $(n+2)$ -hyperspherical coordinates with origin A_0 (\rightarrow 76 Conformal Geometry). Then denoting

by $A \cdot B$ the inner product of hyperspheres A, B , we obtain

$$A_\alpha \cdot A_\beta = g_{\alpha\beta}, \quad \alpha, \beta = 0, 1, \dots, n, \infty,$$

where

$$(g_{\alpha\beta}) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & g_{ij} & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

$$i, j = 1, \dots, n, \quad g_{ij} = g_{ji}.$$

Let z^α be homogeneous coordinates with respect to \mathfrak{R} . Then the Möbius transformation $z \rightarrow \bar{z}$ of S^n is characterized by $\bar{z}^\alpha = c_\beta^\alpha z^\beta$, where $g_{\alpha\beta} c_\alpha^\sigma c_\beta^\tau = g_{\sigma\tau}$, $|c_\beta^\alpha| \neq 0$. The differential of the family of the frames is defined by

$$dA_\alpha = \sum_{\beta=0,1}^{n,\infty} \omega_\alpha^\beta A_\beta, \quad (1)$$

where

$$(\omega_\alpha^\beta) = \begin{pmatrix} \omega_0^0 & \omega_0^j & 0 \\ \sum g_{ik} \omega_k^\infty & \omega_i^j & \sum g_{ik} \omega_0^k \\ 0 & \omega_\infty & -\omega_0^0 \end{pmatrix},$$

$$\sum (g_{ik} \omega_j^k + g_{jk} \omega_i^k) = dg_{ij}.$$

There are $(n+1)(n+2)/2$ linearly independent forms among ω , and this is the number of parameters of the Möbius transformation group. The structure equations of this group are

$$d\omega_\alpha^\beta = \sum \omega_\alpha^\sigma \wedge \omega_\sigma^\beta. \quad (2)$$

The theme of **conformal differential geometry** is the properties of Pfaffian forms ω_α^β satisfying (1) and (2).

Consider a transformation $c: \sum z^\alpha A_\alpha \rightarrow \sum z^\alpha (A_\alpha + dA_\alpha)$. (1) If all ω vanish except ω_0^0 , then all the circles through A_0, A_∞ are invariant, and any point P is transformed to a neighboring point \bar{P} on the circle, such that the cross ratio $(P, \bar{P}; A_0, A_\infty)$ is constant. This transformation is called the **homothety** with centers A_0, A_∞ . (2) If all ω vanish except ω_0^j , $\omega_i^\infty = \sum g_{ik} \omega_0^k$, then all the circles tangent to a fixed direction at A_∞ are transformed into themselves, and any hypersphere through A_∞ and orthogonal to those circles is transformed into a hypersphere having the same property. This transformation is called the **elation** with center A_∞ . (3) If all ω vanish except $\omega_i^j = \sum g_{ik} \omega_\infty^k, \omega_\infty^j$, then the transformation is an elation with center A_0 . (4) If all ω vanish except ω_i^j , then the transformation is an infinitesimal rotation with center A_0 , with A_∞ regarded as a point at infinity. Thus any infinitesimal Möbius transformation is decomposed into the previous four types of transformation.

To study the theory of curves and hypersurfaces in S^n , we again utilize the Frenet frame chosen from a family of frames associated with

A_0 . For example, the Frenet formula of a curve in S^3 is given by

$$(\omega_\alpha^\beta) = \begin{pmatrix} 0 & d\sigma & 0 & 0 & 0 \\ \kappa d\sigma & 0 & 0 & 0 & d\sigma \\ -d\sigma & 0 & 0 & \tau d\sigma & 0 \\ 0 & 0 & -\tau d\sigma & 0 & 0 \\ 0 & \kappa d\sigma & -d\sigma & 0 & 0 \end{pmatrix}.$$

We call $d\sigma$, κ , and τ the **conformal arc element**, **conformal curvature**, and **conformal torsion**, respectively. There are many results on conformal deformation [5].

Concerning Laguerre differential geometry, we have results dual to those in conformal differential geometry (a point is replaced by a straight line and an angle by a distance between the points of contact of the common tangents of two oriented circles).

E. Contact Manifolds

Consider a $(2n+1)$ -dimensional differentiable manifold M^{2n+1} with a 1-form η such that $\eta \wedge (d\eta)^n \neq 0$, where $d\eta$ is the exterior derivative of η and \wedge denotes exterior multiplication. (Note that this is true for the 1-form in the left-hand side of eq. (2) in 82 Contact Transformations A.) Such a manifold is called a **contact manifold** with **contact form** η . The structure group of the tangent bundle of a contact manifold M^{2n+1} reduces to $U(n) \times 1$, where $U(n)$ is the unitary group; hence every contact manifold is orientable. Simple but typical examples are given by the unit sphere S^{2n+1} in Euclidean space E^{2n+2} and the tangent sphere bundle of an $(n+1)$ -dimensional Riemannian manifold M^{n+1} , both with natural contact forms (S. S. Chern [9]). Every 3-dimensional compact orientable differentiable manifold is a contact manifold (J. Martinet [12]).

Now a differentiable manifold M^{2n+1} is said to be an **almost contact manifold** if it admits a tensor field φ of type $(1, 1)$, a vector field ξ , and a 1-form η such that

$$\varphi^2 X = -X + \eta(X)\xi, \quad \eta(\xi) = 1, \quad (3)$$

where X is an arbitrary vector field on M^{2n+1} ; and the triple (φ, ξ, η) is then called an **almost contact structure**. (3) implies that $\varphi\xi = 0$ and $\eta(\varphi X) = 0$ (S. Sasaki [14, I]). The structure group of the tangent bundle of an almost contact manifold M^{2n+1} reduces to $U(n) \times 1$. Indeed, J. W. Gray [10] took this property as his definition of almost contact structure. For any pair of vector fields X and Y on M^{2n+1} , let

$$\begin{aligned} N(X, Y) = & [X, Y] + \varphi[\varphi X, Y] \\ & + \varphi[X, \varphi Y] - [\varphi X, \varphi Y] \\ & - \{X \cdot \eta(Y) - Y \cdot \eta(X)\}\xi, \end{aligned}$$

where $[,]$ is the Poisson bracket; then N is a tensor field of type $(1, 2)$ over M^{2n+1} , which we call the **torsion tensor** of the almost contact structure (φ, ξ, η) . When N vanishes identically on M^{2n+1} , we say that the almost contact structure is **normal**.

An almost contact structure (φ, ξ, η) on M^{2n+1} induces naturally an almost complex structure J on $M^{2n+1} \times R$ (resp. $M^{2n+1} \times S^1$), which reduces to a complex structure if and only if (φ, ξ, η) is normal. A similar statement is also valid for the product space of two almost contact manifolds (A. Morimoto [13]).

If M^{2n+1} is an almost contact manifold with structure tensor (φ, ξ, η) , we can find a positive definite Riemannian metric g so that $g(\varphi X, \varphi Y) = g(X, Y) - \eta(X)\eta(Y)$ for any pair of vector fields X and Y , and the set (φ, ξ, η, g) is then said to be an **almost contact metric structure**.

When M^{2n+1} is a contact manifold with contact form η , there exists a unique vector field ξ which satisfies $d\eta(X, \xi) = 0$, $\eta(\xi) = 1$ for any vector field X . We can then find a tensor field φ of type $(1, 1)$ and a positive definite metric tensor g so that (i) $d\eta(X, Y) = g(X, \varphi Y)$ is satisfied for any pair of vector fields X and Y , and (ii) (φ, ξ, η, g) is an almost contact metric structure. The almost contact metric structure determined in this way by a contact form η is called a **contact metric structure**. A differentiable manifold with normal contact metric structure is called a **normal contact Riemannian manifold** or a **Sasakian manifold**. Brieskorn manifolds are examples of such manifolds. They include, besides the standard sphere S^{2n+1} , all exotic $(2n+1)$ -spheres that bound compact oriented parallelizable manifolds. An almost contact manifold is said to be **regular** or **nonregular** according as the integral curve of ξ is regular or not as a submanifold. A compact regular contact manifold is a principal circle bundle over a symplectic manifold, and it admits a normal contact metric structure if and only if the base manifold is a Hodge manifold (Boothby and Wang [8], Hatakeyama [11]).

Many research papers on the topology and differential geometry of manifolds with the structures defined above have been published by S. Tanno, S. Tachibana, D. E. Blair, M. Okumura, K. Ogiue, S. I. Goldberg, and others.

References

- [1] W. Blaschke, Vorlesungen über Differentialgeometrie, Springer, II, 1923; III, 1929 (Chelsea, 1967).

- [2] G. Bol, Projektive Differentialgeometrie I, II, Vandenhoeck & Ruprecht, 1950.
- [3] E. Cartan, La théorie des groupes finis et continus, Gauthier-Villars, 1951.
- [4] E. Cartan, Leçons sur la théorie des espaces à connexion projective, Gauthier-Villars, 1937.
- [5] P. C. Delens, Méthodes et problèmes des géométries différentielles euclidienne et conforme, Gauthier-Villars, 1927.
- [6] G. Fubini and E. Čech, Introduction à la géométrie projective différentielle des surfaces, Gauthier-Villars, 1931.
- [7] J. Kanitani, Géométrie différentielle projective des hypersurfaces, Mem. Ryojun Coll. of Eng., 1931.
- [8] W. M. Boothby and H. C. Wang, On contact manifolds, Ann. Math., (2) 68 (1958), 721–734.
- [9] S. S. Chern, Pseudo-groupes continus infinis, Coll. int. du CNRS, Geom. Diff., Strasbourg (1953), 119–135.
- [10] J. W. Gray, Some global properties of contact structure, Ann. Math., (2) 69 (1959), 421–450.
- [11] Y. Hatakeyama, Some notes on differentiable manifolds with almost contact structure, Tôhoku Math. J., (2) 15 (1963), 176–181.
- [12] J. Martinet, Formes de contact sur les variétés de dimension 3, Proc. Liverpool Singularities Symposium II (1971), 142–163.
- [13] A. Morimoto, On normal almost contact structures, J. Math. Soc. Japan, 15 (1963), 420–436.
- [14] S. Sasaki, On differentiable manifolds with certain structures which are closely related to almost contact structure. I, Tôhoku Math. J., (2) 12 (1960), 459–476. II (with Y. Hatakeyama), *ibid.*, 13 (1961), 281–294.
- [15] K. Yano and S. Ishihara, Tangent and cotangent bundles, Dekker, 1973.
- [16] D. E. Blair, Contact manifolds in Riemannian geometry, Lecture notes in math. 509, Springer, 1976.

111 (VII.12) Differential Geometry of Curves and Surfaces

A. General Remarks

Let f be an m -dimensional \dagger differentiable manifold M of class C^r into an n -dimensional Euclidean space E^n . More precisely, f is a differentiable mapping of class C^r such that the \dagger differential df_p is injective at every point p of M . The pair (M, f) is called

an **immersed submanifold** (or a **surface**) of E^n . When $m = 1$, we call it a **curve** of E^n , and when $m = n - 1$, a **hypersurface** in E^n . The cases of $n = 2$ and $n = 3$ have been the main objects of study in differential geometry of curves and surfaces. The differential-geometric properties for the general case of immersion are discussed in 365 Riemannian Submanifolds.

B. Frames in E^n

Every \dagger Euclidean motion in E^n can be expressed as the product of a parallel translation and an \dagger orthogonal transformation that keeps the origin of E^n fixed. The set of all parallel translations is a commutative group that can be identified with \mathbf{R}^n . It is a normal subgroup of the group of motions $I(E^n)$ of E^n . So we see that $I(E^n)$ is a \dagger semidirect product of \mathbf{R}^n and the \dagger orthogonal group $O(n)$. The Lie algebra of $I(E^n)$ is the direct sum of \mathbf{R}^n and the Lie algebra $\mathfrak{o}(n)$ of the orthogonal group, where both are regarded as additive groups. Corresponding to this decomposition, we can write the \dagger Maurer-Cartan differential form over $I(E^n)$ as $\omega + \Omega$, where ω belongs to \mathbf{R}^n and Ω to $\mathfrak{o}(n)$. The \dagger structural equation $d(\omega + \Omega) = -(1/2)(\omega + \Omega) \wedge (\omega + \Omega)$ can be divided into the following two parts: $d\omega = \Omega \wedge \omega$; $d\Omega = -(1/2)\Omega \wedge \Omega$. These are known as the **structure equations** of E^n . By an **orthogonal frame** in E^n we mean an ordered set $(x, \mathbf{e}_1, \dots, \mathbf{e}_n)$ consisting of a point x and a set of \dagger orthonormal vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. We denote by $\mathcal{O}(n)$ the set of all orthogonal frames in E^n . If we denote the translation identified with $x \in \mathbf{R}^n$ by T_x , then there is a one-to-one correspondence $\varphi: I(E^n) \rightarrow \mathcal{O}(n)$ given by $\varphi(T_x A) = (x, A\mathbf{e}_1, \dots, A\mathbf{e}_n)$ ($A \in O(n)$). We can make $\mathcal{O}(n)$ into a differentiable manifold so that φ is a \dagger diffeomorphism. We denote the differential forms over $\mathcal{O}(n)$, which are images of ω and Ω under the \dagger dual mapping of φ^{-1} , by the same letters ω and Ω , respectively. For $\mathcal{O}(n)$ as a \dagger principal fiber bundle over \mathbf{R}^n with the projection $\pi: \pi(x, \mathbf{e}_1, \dots, \mathbf{e}_n) = x$ and n vector-valued functions $\varphi_i: \varphi_i(x, \mathbf{e}_1, \dots, \mathbf{e}_n) = \mathbf{e}_i$ over $\mathcal{O}(n)$, we have

$$\begin{aligned} \omega &= \sum_i \omega^i \mathbf{e}_i, & \Omega &= \sum_{i < j} \Omega^{ij} E_{ij}, \\ \omega^i &= (d\pi, \varphi_i), & \Omega^{ij} &= (d\varphi_i, \varphi_j), \end{aligned} \quad (1)$$

$$d\omega^i = \sum_j \Omega^{ij} \wedge \omega^j, \quad d\Omega^{ij} = \sum_k \Omega^{ik} \wedge \Omega^{kj},$$

where $\{E_{ij}\}$ is a basis of $\mathfrak{o}(n)$ defined by $E_{ij}\mathbf{e}_j = \mathbf{e}_i$, $E_{ij}\mathbf{e}_i = -\mathbf{e}_j$, $E_{ij}\mathbf{e}_k = 0$ ($k \neq i, j$) and $(\ , \)$ is the scalar product of vector-valued forms induced from the scalar product of E^n . Any diffeomorphism of $\mathcal{O}(n)$ onto itself preserving ω and Ω must be a Euclidean motion.

C. Theory of Curves

Let (M, f) be an immersion of a 1-dimensional differentiable manifold M into E^n . We identify the tangent space of E^n at each point with E^n itself. Then df_x maps the origin of the tangent space M_x to $f(x)$, and the image $df_x(M_x)$ of M_x by df_x is a straight line passing through $f(x)$ in E^n , called the **tangent line** of $f(M)$ at $f(x)$. By $\mathcal{O}_f(M)$ we mean the set of all ordered sets $(x, \mathbf{e}_1, \dots, \mathbf{e}_n)$, where $x \in M$ and $\{\mathbf{e}_i\}$ is an orthonormal basis of E^n such that $\mathbf{e}_1 \in df_x(M_x)$. Then $\mathcal{O}_f(M)$ can be naturally immersed in $\mathcal{O}(n)$ by the mapping $\hat{f}: \hat{f}(x, \mathbf{e}_1, \dots, \mathbf{e}_n) = (f(x), \mathbf{e}_1, \dots, \mathbf{e}_n)$. We can pull back the differential forms $\omega, \Omega, \omega', \Omega^i$ over $\mathcal{O}(n)$ to $\mathcal{O}_f(M)$ \hat{f}^* and denote them by θ, Θ, θ^i , and Θ^i , respectively; then we have $\theta^i = 0$ ($i > 1$). Let f_1 and f_2 be two immersions of M into E^n . Then in order for there to exist a Euclidean motion α of E^n such that $f_1 = \alpha \circ f_2$, it is necessary and sufficient that there exist a diffeomorphism φ of $\mathcal{O}_{f_1}(M)$ onto $\mathcal{O}_{f_2}(M)$ such that $\theta_{f_1} = \varphi^*(\theta_{f_2})$, $\Theta_{f_1} = \varphi^*(\Theta_{f_2})$. Let π_f be the projection of the fiber bundle $\mathcal{O}_f(M)$, and let φ_i be naturally defined vector-valued functions over $\mathcal{O}_f(M)$. Then we have $d(f \circ \pi_f) = \theta^1 \varphi_1$, $d\varphi_i = \sum_{j=1}^n \Theta^{ij} \varphi_j$. If we put $ds^2(X) = \|df_x(X)\|^2$ ($X \in M_x$), then we have $(\theta^1)^2 = \pi_f^*(ds^2)$. For each point $x \in M$ there are two possibilities for the choice of \mathbf{e}_1 corresponding to two orientations of the curve. But since $(d\varphi_1, d\varphi_1) = \pi_f^*(\rho^2 ds^2)$, ρ^2 depends only on the point x of M . We call ρ (≥ 0) the **absolute curvature**. We now choose an orientation of the curve and then \mathbf{e}_1 in accordance with the orientation. Thus we get a submanifold of $\mathcal{O}_f(M)$, which we again express by the notation $\mathcal{O}_f(M)$. If we define the form ds by $ds(X) = (df(X), \mathbf{e}_1)$, we have $\theta^1 = \pi_f^*(ds)$, and ds is called the **line element**. Any t local cross section $R: \pi_f \circ R = 1$ of the bundle $\mathcal{O}_f(M)$ is called a **moving frame**. Putting

$$R^*(\theta^1) = ds, \quad R^*(\Theta^i) = \rho^i ds,$$

we see that the following equation holds over M :

$$d\mathbf{e}_i = \sum_j \rho^{ij} ds \mathbf{e}_j. \quad (2)$$

For two immersions (M, f_1) and (M, f_2) , we have $f_1 = \alpha \circ f_2$ (α is a Euclidean motion) if and only if they have the same ds and ρ^{ij} for some moving frames.

D. Frenet's Formulas

In order to study local properties of curves it is sufficient to consider them on t Jordan arcs of class C^r . With respect to orthogonal coordinates (x^1, \dots, x^n) in E^n , such a curve is represented parametrically by $x^i = f^i(t)$ ($t \in [a, b]$, $\sum (dx^i/dt)^2 > 0$) or by a vector representation

$\mathbf{x} = \mathbf{x}(t)$. If φ is a diffeomorphism of a closed interval $[a', b']$ onto $[a, b]$, then $f \circ \varphi$ and f are representations of the same arc in E^n , and φ is called a **transformation of the parameter**. Any curve of class C^1 is t rectifiable, and its **arc length** is given by $s = \int_a^b (\sum_{i=1}^n (dx^i/dt)^2)^{1/2} dt$. We may choose the arc length s measured from a point on the arc as a parameter, called the **canonical parameter** of the arc. Consider an arc C of class C^n given by the vector representation $\mathbf{x} = \mathbf{x}(s)$, $s \in [a, b]$. We assume that its t Wronskian $|\mathbf{x}'(s), \dots, \mathbf{x}^{(n)}(s)|$ is not identically zero (we denote by $'$ the derivative with respect to the canonical parameter), which means that the arc C is not contained in a hyperplane in E^n . A point at which the Wronskian vanishes is called a **stationary point**, and we assume that there exists no stationary point on C . By the t Gram-Schmidt orthonormalizing process we obtain an orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ ($|\mathbf{e}_1, \dots, \mathbf{e}_n| > 0$) from n vectors $\mathbf{x}'(s), \dots, \mathbf{x}^{(n)}(s)$ at each point of C . We call the frame thus determined the **Frenet frame**. With respect to the Frenet frame, (2) is rewritten as

$$\begin{aligned} \mathbf{e}_i'(s) &= -\kappa_{i-1}(s)\mathbf{e}_{i-1}(s) + \kappa_i(s)\mathbf{e}_{i+1}(s), \\ i &= 1, \dots, n; \\ \kappa_0(s) &= \kappa_n(s) = 0; \\ \kappa_j(s) &> 0, \quad j &= 1, \dots, n-2. \end{aligned} \quad (3)$$

These are called **Frenet's formulas** (or the **Frenet-Serret formulas**). We call $\kappa_1, \kappa_2, \dots, \kappa_{n-2}$ the first, second, \dots , $(n-2)$ nd **curvature**, respectively, while we call κ_{n-1} the **torsion** for $n \geq 3$. For a curve in a lower-dimensional subspace $E^m \subset E^n$, we set $\kappa_i = 0$ ($i > m$). The curvatures and the torsion of a straight line are zero. To get Frenet's formulas in these special cases, we fix \mathbf{e}_i ($i > m$) in the subspace orthogonally complementary to E^m in E^n and proceed as in the general case. Suppose that C_1, C_2 are arcs such that both of their Frenet frames are of class C^1 . If there exists a diffeomorphism of C_1 to C_2 that preserves arc length and the κ_i ($i = 1, \dots, n-1$) are equal at corresponding points, C_1 and C_2 are mapped onto each other by a motion of E^n . This is the **fundamental theorem of the theory of curves**.

Given $n-1$ functions of class C^1 $\kappa_1(s) \geq 0, \dots, \kappa_{n-2}(s) \geq 0$ (we assume that the equality signs occur at most at a finite number of points) and $\kappa_{n-1}(s)$ for $0 \leq s \leq L$, there exists an arc that has $\kappa_1, \dots, \kappa_{n-2}, \kappa_{n-1}$ as its first, \dots , $(n-2)$ nd curvatures and its torsion, respectively. The equations $\kappa_i = \kappa_i(s)$ are called the **natural equations** of the curve.

E. Plane Curves

Let $\mathbf{x} = \mathbf{x}(s)$ be a curve of class C^2 in E^2 , and $(\mathbf{x}(s), \mathbf{e}_1, \mathbf{e}_2)$ its Frenet frame. The tangent and

the normal of this curve at $\mathbf{x}(s)$ have parametric representations $\mathbf{x}(s) + t\mathbf{e}_1$, $\mathbf{x}(s) + t\mathbf{e}_2$, respectively (with parameter t). Frenet's formulas are written as $\mathbf{x}' = \mathbf{e}_1$, $\mathbf{x}'' = \mathbf{e}_1' = \kappa\mathbf{e}_2$, and ρ is called the **curvature** of the curve C . The natural equation is given by $\kappa = \kappa(s)$. If $\kappa(s) = \text{constant} \neq 0$ along C , C must be a portion of a circle. Another way of defining the curvature is as follows: We take a fixed direction (for example, the positive direction of the x -axis on E^2) and denote by $\theta(s)$ the angle made by the tangent T_s of the curve C at $\mathbf{x}(s)$ with the direction. Then we have $\kappa(s) = d\theta/ds$. If $n=2$, the curvature can take both positive and negative values. Figs. 1 and 2 suggest a geometric meaning of $\kappa > 0$ and $\kappa < 0$, respectively. The circle with center $\mathbf{x} = \mathbf{x}(s) + (1/\kappa)\mathbf{e}_2$ and radius $1/\kappa$ has a contact of higher order than any other circle in E^2 . We call this circle the **osculating circle** (or **circle of curvature**) at the point $\mathbf{x} = \mathbf{x}(s)$, its center the **center of curvature**, and $1/\kappa$ the **radius of curvature**. The locus C' of the center of curvature of a curve C is called an **evolute** of C . Conversely, C is called an **involute** of C' , the envelope of the family of normal lines of C . When a curve is given in terms of its canonical parameter s , the curvature is given by $|\mathbf{x}'(s), \mathbf{x}''(s)|$; when the curve is given by another parameter t as $\mathbf{x} = \mathbf{x}(t)$, the curvature is given by $\kappa(t) = |\mathbf{x}'(t), \mathbf{x}''(t)|/|\mathbf{x}'(t)|^3$, where $'$ means d/dt .

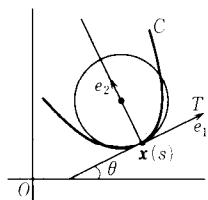


Fig. 1
 $\kappa > 0$.

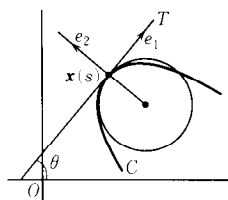


Fig. 2
 $\kappa < 0$.

The facts we have just stated concern local properties of plane curves. We shall now discuss the global theory of curves, which deals with properties of each curve as a whole. Let $C: \mathbf{x} = \mathbf{x}(s)$, $a \leq s \leq b$, be a closed curve. Let $\theta(s)$, $0 \leq \theta(s) < 2\pi$, be the angle that $\mathbf{e}_1(s)$ makes with the x axis. Put $\Theta = \int_a^b \theta'(s) ds$. Intuitively, Θ measures the total rotation of $\mathbf{e}_1(s)$ as we run along the curve C from a to b . Since C is closed, Θ is an integer multiple I of 2π . The integer I is called the **rotation number** of C , and is equal to $(1/2\pi) \int_a^b \kappa(s) ds$. Let \bar{D} be a closed domain consisting of points in the interior and on the boundary of a simple closed curve C . C is called a **closed convex curve** or an **oval** if \bar{D} is convex in E^2 . Among all ovals of given length, the circle has the maximum area. Various generalizations of this theorem have been obtained, and the collection of

problems of this kind is called the **isoperimetric problem**. This problem has intimate connections with fields such as integral geometry. The oval has a convenient parameter other than the arc length parameter s . Given a number t , $0 \leq t \leq 2\pi$, there exists a unique point $\mathbf{x}(t)$ in the oval such that $\mathbf{e}_2 = (\cos t, \sin t)$ at $\mathbf{x}(t)$. When we describe the oval in terms of the parameter t , the tangent vector at $\mathbf{x}(t)$ is parallel to that at $\mathbf{x}(t + \pi)$, and we can define the width $W(t)$ at $\mathbf{x}(t)$. $W(t)$ is called the **width of the oval**. A curve is called a **curve of constant width** if the curve is an oval whose width $W(t)$ does not depend on t . The circle is a typical example of a curve of constant width. Reuleaux's triangle is another well-known example of a curve of constant width (Fig. 3). For a curve of constant width of width W and length L , we have $L = \pi W$.

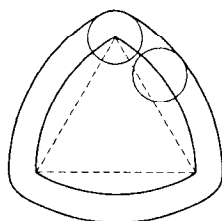


Fig. 3

There are also some results concerning the relations between local properties (for example, curvature) and properties of the whole figure. An example is given by the **four-vertex theorem**. A vertex on a curve C is by definition a point where $d\kappa/ds = 0$. Then there are at least four vertices on an oval of class C^3 . A simple closed curve with $\kappa \geq 0$ (≤ 0) must be convex (\rightarrow 89 Convex Sets).

F. Space Curves

Let $\mathbf{x} = \mathbf{x}(s)$ ($s \in [a, b]$) be a curve C of class C^3 in E^3 defined in terms of the canonical parameter s . Let $(\mathbf{x}(s), \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be Frenet frames along C . Then we have the Frenet formulas

$$\mathbf{e}_1' = \kappa_1 \mathbf{e}_2, \quad \mathbf{e}_2' = -\kappa_1 \mathbf{e}_1 + \kappa_2 \mathbf{e}_3, \quad \mathbf{e}_3' = -\kappa_2 \mathbf{e}_2.$$

We call $1/\kappa_1$, $1/\kappa_2$ the **radius of curvature** and the **radius of torsion**, respectively. The line $\mathbf{x} = \mathbf{x}(s_0) + t\mathbf{e}_1$ is the tangent of C at $\mathbf{x}(s_0)$. The two straight lines through the point $\mathbf{x}(s_0)$ defined by $\bar{\mathbf{x}} = \mathbf{x}(s_0) + t\mathbf{e}_2$ and $\bar{\mathbf{x}} = \mathbf{x}(s_0) + t\mathbf{e}_3$ are called the **principal normal** and the **binormal** of C at $\mathbf{x}(s_0)$, respectively. The three planes through $\mathbf{x}(s_0)$ defined by $\bar{\mathbf{x}} = \mathbf{x}(s_0) + t\mathbf{e}_2 + t\mathbf{e}_3$, $\bar{\mathbf{x}} = \mathbf{x}(s_0) + t\mathbf{e}_3 + t\mathbf{e}_1$, and $\bar{\mathbf{x}} = \mathbf{x}(s_0) + t\mathbf{e}_1 + t\mathbf{e}_2$ are called the **normal plane**, the **rectifying plane**, and the **osculating plane**, respectively.

At a point $\mathbf{x}(s_0)$ of a curve $\mathbf{x} = \mathbf{x}(s)$ of class C^ω , we take $\mathbf{e}_1(s_0)$, $\mathbf{e}_2(s_0)$, and $\mathbf{e}_3(s_0)$ as unit vectors of the coordinate axes. Substituting the Frenet formulas into the Taylor expansion of $\mathbf{x}(s)$, we see that the new coordinate $\mathbf{x}_1(s)$, $\mathbf{x}_2(s)$, $\mathbf{x}_3(s)$ of C are given by

$$x_1 = (s - s_0) - (\kappa_1(s_0)/6)(s - s_0)^3 + \dots,$$

$$x_2 = (\kappa_1(s_0)/2)(s - s_0)^2 + (\kappa'_1(s_0)/6)(s - s_0)^3 + \dots,$$

$$x_3 = (\kappa_1(s_0)\kappa_2(s_0)/6)(s - s_0)^3 + \dots$$

These are called **Bouquet's formulas**. Utilizing these formulas we can see the nature of the curve with given κ_1 and κ_2 . A curve and its osculating plane at a point on it have contact of order higher than any other plane through that point. The family of osculating planes of C envelopes a developable surface S and coincides with the locus of tangent lines to C . We call S the **tangent surface** of C , and C the **line of regression** of S . The family of rectifying planes of C also envelopes a developable surface called the **rectifying surface**, and C is a geodesic on this surface. The family of normal planes of C envelopes either a cone or a tangent surface of another curve \bar{C} . When the natural equation of a space curve has a special form, the shape of the curve is simple. For example, $\kappa_1(s) = \text{constant}$, $\kappa_2(s) = \text{constant}$ represent a curve, called an **ordinary helix**, on a cylinder which cuts all the generators of the cylinder at a constant angle. More generally, it is known that if $\kappa_1/\kappa_2 = \text{constant}$, the tangent at each point of the curve makes a constant angle with a fixed direction. Such a curve is called a **generalized helix** or a **curve of constant inclination**. Each curve satisfying $ak_1 + bk_2 = c$ ($ab \neq 0$) is called a **Bertrand curve**. For a Bertrand curve there exists another curve \bar{C} and a correspondence of C onto \bar{C} such that they have a common principal normal at corresponding points. Conversely, this property is also a sufficient condition for C to be a Bertrand curve. A **Mannheim curve** is defined analogously as a curve having a correspondence with another curve \bar{C} such that the principal normal of C and the binormal of \bar{C} coincide at corresponding points. When a correspondence of C and \bar{C} has the property that tangents at corresponding points are parallel, then the correspondence is called a **correspondence of Combesure**.

We have stated mainly local properties of space curves. There are also several results about global properties of curves in E^3 analogous to the case of plane curves. For a simple closed curve C of length L , we call $K = \int_0^L \kappa_1(s) ds$ the **total curvature** of C . Generally we have $K \leq 2\pi$, while $K = 2\pi$ if and only if C is a closed convex curve lying in a plane (W. Fenchel) [5, 6]. The total curvature is deeply

related to the properties of knots. If a simple closed curve in E^3 is knotted, then the total curvature is at least 4π [7, 8]. We fix an origin O in E^3 and draw a unit tangent vector with initial point O parallel to the unit tangent vector at each point of a space curve C ; then the endpoint of this vector traces a curve \bar{C} on the unit sphere with the center O . We call \bar{C} the **spherical indicatrix** of C and the correspondence of C to \bar{C} a **spherical representation**. The total curvature K of a curve C is equal to the length of \bar{C} . Consequently, we have $K = \oint_C d\theta$, where θ is the angular deflection of the tangent line along the closed curve C .

G. Theory of Hypersurfaces

Let (M, f) be an immersion of an $(n-1)$ -dimensional differentiable manifold M of class C^r into E^n . Then we can define on the hypersurface M a positive definite differential form g of degree 2 induced from the inner product of E^n : $g_x(X, X) = (df_x(X), df_x(X))$, $X \in M_x$. Then M becomes a Riemannian manifold with Riemannian metric g . We call g the **first fundamental form** of (M, f) . The Riemannian geometry on a surface with its first fundamental form as Riemannian metric is called **geometry on a surface** (\rightarrow 364 Riemannian Manifolds).

By $\mathcal{O}_f(M)$ we mean the set of all the ordered sets $(x, \mathbf{e}_1, \dots, \mathbf{e}_n) \in \mathcal{O}(n)$, where $x \in M$ and $\{\mathbf{e}_i\}$ ($i = 1, 2, \dots, n$) is an orthonormal system of E^n such that $\mathbf{e}_i \in df_x(M_x)$ ($i = 1, \dots, n-1$). Then $\mathcal{O}_f(M)$ with natural projection π_f and natural differentiable structure is a principal fiber bundle over M and has a natural immersion $\hat{f}: \hat{f}(x, \mathbf{e}_1, \dots, \mathbf{e}_n) = (f(x), \mathbf{e}_1, \dots, \mathbf{e}_n)$ in the principal fiber bundle $\mathcal{O}(n)$. We can pull back the forms on $\mathcal{O}(n)$ to $\mathcal{O}_f(M)$ by \hat{f}^* and put $\theta = \hat{f}^*(\omega)$, $\Theta = \hat{f}^*(\Omega)$; then the structural equations of E^n are transformed to $d\theta = \Theta \wedge \theta$, $d\Theta = (-1/2)\Theta \wedge \Theta$. Furthermore, if we put $\theta^i = \hat{f}^*(\omega^i)$, $\Theta^{ij} = \hat{f}^*(\Omega^{ij})$, then $\theta^n = 0$ and θ^i , Θ^{ij} ($i, j < n$) depend only on the first fundamental form of (M, f) . Let f_1 and f_2 be two immersions of M into E^n . Then in order that there exist a Euclidean motion α of E^n such that $f_1 = \alpha \circ f_2$, it is necessary and sufficient that there exist a diffeomorphism φ of $\mathcal{O}_{f_1}(M)$ onto $\mathcal{O}_{f_2}(M)$ such that $\theta_{f_1} = \varphi^*(\theta_{f_2})$ and $\Theta_{f_1} = \varphi^*(\Theta_{f_2})$. Suppose that M is orientable and oriented. Then the unit vector field normal to $df_x(M_x)$ at every point $x \in M$ in E^n defines a mapping of M into the unit sphere in E^n called the **spherical representation** of M or the **Gauss mapping** (**Gauss map**). Regarding the unit normal vector field ξ of M as a vector-valued function over M , we can define a symmetric product of df and $d\xi$ by $-(df, d\xi)(X, Y) =$

$(1/2)[(df(X), d\xi(Y)) + (df(Y), d\xi(X))]$, called the **second fundamental form** of (M, f) .

Two immersions f_1 and f_2 of M that induce the same first and second fundamental forms have a Euclidean motion α such that $f_1 = \alpha \circ f_2$; and the converse is also true. This fact is called the **fundamental theorem of the theory of surfaces**.

H. Theory of Surfaces in E^3 (\rightarrow 365 Riemannian Submanifolds; Appendix A, Table 4.I)

A surface in E^n is locally expressed by parametric equations $x_i = x_i(u_\alpha)$ ($i = 1, \dots, n$; $\alpha = 1, \dots, m$) or by a single vector equation $\mathbf{x} = \mathbf{x}(u_\alpha)$. We are mainly concerned with the case $n = 3$, $m = 2$, and we express the surface by a vector representation $\mathbf{x} = \mathbf{x}(u, v)$. The first and second fundamental forms are written as

$$E du^2 + 2F du dv + G dv^2,$$

$$L du^2 + 2M du dv + N dv^2.$$

If we use the usual notation of tensor analysis, then the first and second fundamental forms are also denoted by $g_{\alpha\beta} du^\alpha du^\beta$ and $H_{\alpha\beta} du^\alpha du^\beta$, respectively, where u^α ($\alpha = 1, 2$) are parameters (with Σ omitted by Einstein's convention). We call $\{g_{\alpha\beta}\}$, $\{H_{\alpha\beta}\}$ the **first** and **second fundamental quantities**, respectively (\rightarrow 365 Riemannian Submanifolds). At the point $p_0 = \mathbf{x}(u_0, v_0)$ on a surface S that corresponds to parameter values (u_0, v_0) , the curves expressed by $v = v_0$ and $u = u_0$ are called a ***u*-curve** and a ***v*-curve** through p_0 , respectively. Let $\mathbf{x}_u, \mathbf{x}_v$ denote the tangent vectors $\partial\mathbf{x}/\partial u, \partial\mathbf{x}/\partial v$ at p_0 to the *u*-curve and *v*-curve, respectively, through the given point p_0 and ξ denote the unit vector orthogonal to \mathbf{x}_u and \mathbf{x}_v . Then ξ is called the **normal vector** of S at p_0 and $(\mathbf{x}_u, \mathbf{x}_v, \xi)$ the **Gaussian frame** of S at p_0 . Although a Gaussian frame is not in general an orthogonal frame, it is intimately related to local parameters. The plane that passes through the point p_0 and is spanned by $\mathbf{x}_u, \mathbf{x}_v$ is called the **tangent plane** to S at p_0 . The coefficients of the second fundamental form $L(u, v)$, $M(u, v)$, $N(u, v)$ are expressed by the inner products $L = (-\mathbf{x}_u, \xi_u)$, $M = (-\mathbf{x}_u, \xi_v)$, $N = (-\mathbf{x}_v, \xi_v)$ ($\xi_u = \partial\xi/\partial u$, $\xi_v = \partial\xi/\partial v$).

Let (X, Y) be the coordinates of a point on the tangent space at p_0 with respect to the Gaussian frame. We call the curve of the second order defined by $LX^2 + 2MXY + NY^2 = \varepsilon$ (ε is a suitable constant) the **Dupin indicatrix**. The point p_0 is called an **elliptic point** or a **hyperbolic point** on S according as the Dupin indicatrix at the point is an ellipse or a hyperbola. If p_0 is an elliptic point, then points near p_0 on the surface lie on one side of the tangent plane at p_0 , whereas if p_0 is a hyperbolic point,

points near p_0 on the surface lie on both sides of the tangent plane at p_0 (Figs. 4, 5). A hyperbolic point is also called a **saddle point**, since in a neighborhood of the point the surface looks like a saddle. A point that is neither elliptic nor hyperbolic is called a **parabolic point**; at a parabolic point we have $LN - M^2 = 0$. If at least one of L, M, N does not vanish at p_0 , then there is a neighborhood of p_0 of the surface that lies on one side of the tangent plane at p_0 (Fig. 6). If a vector (X, Y) on the tangent plane of the surface at p_0 satisfies the equation $LX^2 + 2MXY + NY^2 = 0$, then the direction of the vector is called an **asymptotic direction**. If the point p_0 is elliptic, such a direction does not exist; if p_0 is hyperbolic, the direction is an asymptotic direction of the Dupin indicatrix on the tangent plane at p_0 . A curve C on a surface such that the tangent line at each point of the curve coincides with an asymptotic direction of the surface at the point is called an **asymptotic curve**.

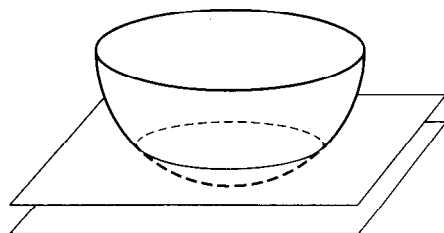


Fig. 4
Elliptic point.

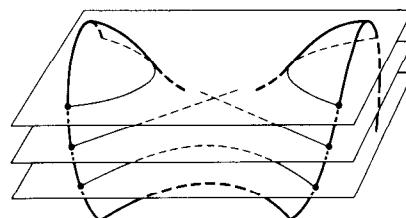


Fig. 5
Hyperbolic point.

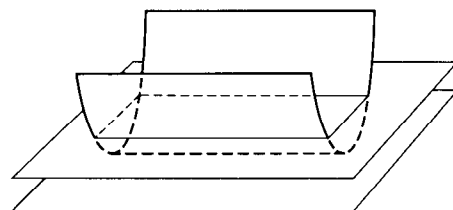


Fig. 6
Parabolic point.

Let $C: \mathbf{x} = \mathbf{x}(u(t), v(t))$ be a curve through p_0 on the surface $\mathbf{x} = \mathbf{x}(u, v)$. Then the curvature κ of C as a space curve is given by

$$\kappa \cos \theta = \frac{L du^2 + 2M du dv + N dv^2}{E du^2 + 2F du dv + G dv^2},$$

where $du dv$ is the direction of C on the surface at p_0 and θ is the angle between the normal of the surface at p_0 and the principal normal of C at p_0 . The center of curvature at a point p_0 of a curve C of class C^2 on a surface of class C^2 is the projection on its osculating plane of the center of curvature of the section C^* (of the surface) cut by the plane determined by the tangent to the curve at p_0 and the normal of the surface at the point (**Meusnier's theorem**). The curvature of the curve C^* at p_0 is called the **normal curvature** of the surface at the point for the tangent direction. Since the normal curvature for a direction at a point is a continuous function of this direction that can be represented as a point on a unit circle, there exist two directions that realize the maximum and minimum of the normal curvature. These directions are given by the equation

$$\begin{vmatrix} E du + F dv & F du + G dv \\ L du + M dv & M du + N dv \end{vmatrix} = 0.$$

When this quadratic equation in du/dv has nonzero discriminant, it determines two directions defined by its two roots. These directions are called **principal directions** at the point. A curve C on a surface such that the tangent line at each point of the curve coincides with a principal direction at the point is called a **line of curvature**. When all lines of curvature of a surface are circles, the surface is called a **cyclide of Dupin**. The two normal curvatures corresponding to two principal directions are given by $1/R$, satisfying the following second-order equation:

$$\left(\frac{1}{R}\right)^2 - \frac{EN + GL - 2FM}{EG - F^2} \frac{1}{R} + \frac{LN - M^2}{EG - F^2} = 0.$$

They are called **principal curvatures**, and each of their inverses is called a **radius of principal curvature**. The mean value $H = (\kappa_1 + \kappa_2)/2$ of two principal curvatures $\kappa_i = 1/R_i$ ($i = 1, 2$) is called the **mean curvature** (or **Germain's curvature**), and the product $K = \kappa_1 \kappa_2$ is called the **total curvature** (or **Gaussian curvature**). These are given by

$$H = \frac{1}{2} \frac{EN + GL - 2FM}{EG - F^2}, \quad K = \frac{LN - M^2}{EG - F^2}.$$

A point on a surface is elliptic, hyperbolic, or parabolic according as $K > 0$, $K < 0$, or $K = 0$ at the point. A point where the second fundamental form is proportional to the first fundamental form is called an **umbilical point**, and a point where the second fundamental form vanishes is called a **flat point** or a **geodesic point**. If a surface consists of umbilical points only, the ratio $(L(du)^2 + 2M du dv + N(dv)^2)/(E(du)^2 + 2F du dv + G(dv)^2)$ is a constant, and the surface is either a sphere or a portion of it. If a surface consists of flat points

only, the surface must be either a plane or a portion of it. The mean curvature and the Gaussian curvature of a sphere are constant, and those of a plane are both equal to zero. If we use the spherical representation of a surface stated in Section G, we can give to the Gaussian curvature the following geometric meaning: Let A be the area of the domain enclosed by a closed curve C around a point p_0 on a surface, and let A^* be the area of the domain on the unit sphere enclosed by the curve that is the image of C under the spherical representation of the surface. Then the limit of A^*/A as the closed curve C tends to the point p_0 is equal to K at p_0 .

Let us denote by $(g^{\alpha\beta})$ the inverse matrix of the matrix $(g_{\alpha\beta})$ whose elements are coefficients of the first fundamental form $g_{\alpha\beta} du^\alpha du^\beta$. We easily see that $g^{11} = G/(EG - F^2)$, $g^{12} = g^{21} = -F/(EG - F^2)$, $g^{22} = E/(EG - F^2)$. We introduce the symbols

$$[\beta\gamma, \alpha] = \frac{1}{2} \left(\frac{\partial g_{\alpha\beta}}{\partial u^\gamma} + \frac{\partial g_{\gamma\alpha}}{\partial u^\beta} - \frac{\partial g_{\beta\gamma}}{\partial u^\alpha} \right),$$

$$\begin{Bmatrix} \alpha \\ \beta\gamma \end{Bmatrix} = g^{\alpha\delta} [\beta\gamma, \delta],$$

which are called the **Christoffel symbols** of the first and second kinds, respectively. Suppose that a surface is given by the vector representation $\mathbf{x} = \mathbf{x}(u_1, u_2)$, and put $\mathbf{x}_\alpha = \partial \mathbf{x} / \partial u^\alpha$, $\mathbf{x}_{\alpha\beta} = \partial \mathbf{x}_\alpha / \partial u^\beta$. Then for the derivatives of the Gauss frame, we obtain

$$\mathbf{x}_{\alpha\beta} = \begin{Bmatrix} \gamma \\ \alpha\beta \end{Bmatrix} \mathbf{x}_\gamma + H_{\alpha\beta} \boldsymbol{\xi}, \quad \boldsymbol{\xi}_\alpha = -g^{\gamma\beta} H_{\beta\alpha} \mathbf{x}_\gamma.$$

We call the former **Gauss's formula** and the latter **Weingarten's formula**. The integrability conditions of these partial differential equations are

$$R^\alpha_{\beta\gamma\delta} = H_{\delta\beta} H_\gamma^\alpha - H_{\gamma\beta} H_\delta^\alpha, \quad H_\gamma^\alpha = g^{\alpha\sigma} H_{\gamma\sigma},$$

$$\frac{\partial H_{\alpha\beta}}{\partial u^\gamma} - \frac{\partial H_{\alpha\gamma}}{\partial u^\beta} + \begin{Bmatrix} \sigma \\ \alpha\beta \end{Bmatrix} H_{\sigma\gamma} - \begin{Bmatrix} \sigma \\ \alpha\gamma \end{Bmatrix} H_{\sigma\beta} = 0,$$

where

$$R^\alpha_{\beta\gamma\delta} = \frac{\partial}{\partial u^\gamma} \begin{Bmatrix} \alpha \\ \delta\beta \end{Bmatrix} - \frac{\partial}{\partial u^\delta} \begin{Bmatrix} \alpha \\ \gamma\beta \end{Bmatrix} + \begin{Bmatrix} \sigma \\ \delta\beta \end{Bmatrix} \begin{Bmatrix} \alpha \\ \gamma\sigma \end{Bmatrix} - \begin{Bmatrix} \sigma \\ \gamma\beta \end{Bmatrix} \begin{Bmatrix} \alpha \\ \delta\sigma \end{Bmatrix}$$

are components of the curvature tensor. The former are called the **Gauss equations**, and the latter the **Codazzi-Mainardi equations**. In connection with these equations, **Bonnet's fundamental theorem** states the following: Suppose that a positive definite symmetric matrix $(g_{\alpha\beta})$ and a symmetric matrix $(H_{\alpha\beta})$ are given that are functions of class C^2 and C^1 , respectively, defined over a simply connected domain D in \mathbf{R}^2 . If they satisfy the Gauss

equations and the Codazzi-Mainardi equations, then there exists a surface $\mathbf{x} = \mathbf{x}(u_1, u_2)$ with the given $(g_{\alpha\beta})$ and $(H_{\alpha\beta})$ as coefficients of its first and second fundamental forms, respectively. Such a surface is determined uniquely if, for an arbitrary fixed point (u_1^0, u_2^0) of D , we assign an arbitrary point p_0 and a frame $(\mathbf{x}_1^0, \mathbf{x}_2^0, \xi^0)$ at p_0 so that $\mathbf{x}_1^0, \mathbf{x}_2^0$ are orthogonal to the unit vector ξ^0 and $(\mathbf{x}_\alpha^0, \mathbf{x}_\beta^0) = g_{\alpha\beta}(u_1^0, u_2^0)$ as the Gaussian frame at p_0 . On the other hand, we can investigate surfaces as Riemannian manifolds defined by the first fundamental form.

A \dagger diffeomorphism between two surfaces preserving arc length is called an **isometric mapping**. The condition of preserving arc length is equivalent to the condition that the first fundamental quantities of the surfaces coincide at each pair of corresponding points, provided that we have introduced parameters on the two surfaces so that corresponding points have the same parameter values. In such a case two surfaces are said to be **isometric**. From the Gauss equation we can see that the total curvature depends only on the first fundamental quantities. So K is a quantity that is preserved under isometric mappings (**Gauss's theorema egregium**).

A vector field $\lambda^\alpha(t)\mathbf{x}_\alpha$ defined along a curve $u^\alpha = u^\alpha(t)$ on a surface is said to be **parallel in the sense of Levi-Civita** along the curve if its \dagger covariant derivative along the curve vanishes, i.e., if

$$d\lambda^\alpha/dt = d\lambda^\alpha/dt + \left\{ \begin{matrix} \alpha \\ \beta\gamma \end{matrix} \right\} \lambda^\beta du^\gamma/dt = 0.$$

The length of a vector belonging to a vector field that is parallel along a curve C is constant along C . The angle of two vectors both belonging to vector fields that are parallel along C is also constant along C . Choose two vector fields $\lambda_{(a)}^\alpha$ parallel along a curve C on a surface that satisfy $g_{\alpha\beta}\lambda_{(a)}^\alpha\lambda_{(b)}^\beta = \delta_{ab}$. Then the tangent vector to C is expressed by $du^\alpha/dt = \lambda_{(a)}^\alpha v^a(t)$. Take a 2-plane and fix an orthogonal coordinate system on it; then the integral curve C of a set of ordinary differential equations $dx^\alpha/dt = C_{(a)}^\alpha v^a(t)$ ($C_{(a)}^\alpha = \lambda_{(a)}^\alpha(P_0)$) is called the **development** of C (\rightarrow 80 Connections). We denote by κ the curvature of a curve C of class C^2 on a surface S of class C^2 and by σ the angle between the binormal of C and the normal of S at the same point. Then the quantity $\kappa_g = \kappa \cos \sigma$, belonging to the geometry on the surface, is called the **geodesic curvature** of the curve at the point. A curve with vanishing geodesic curvature is called a **geodesic**. It satisfies the differential equations

$$\frac{d^2 u^\alpha}{ds^2} + \left\{ \begin{matrix} \alpha \\ \beta\gamma \end{matrix} \right\} \frac{du^\beta}{ds} \frac{du^\gamma}{ds} = 0.$$

The development of a geodesic is a straight line (\rightarrow 364 Riemannian Manifolds).

Let us consider a simply connected, orientable bounded domain D on a surface such that the boundary of D is a simple closed curve C that consists of a finite number of arcs of class C^2 . If we denote by α_i ($i = 1, 2, \dots, m$) the external angles at vertices of the curvilinear polygon C (Fig. 7), we have

$$\int_C \kappa_g ds + \sum_{i=1}^m \alpha_i + \iint_D K d\sigma = 2\pi.$$

This is called the **Gauss-Bonnet formula**. In particular, if all the arcs of C are geodesics, we have

$$\sum_{i=1}^m \alpha_i + \iint_D K d\sigma = 2\pi.$$

This formula implies as special cases the following well-known theorems in Euclidean geometry and spherical trigonometry: (i) The sum of interior angles of a triangle is equal to π . (ii) The area of a spherical triangle is proportional to its spherical excess. The formula also implies the following theorem: On any closed orientable surface we have $\iint K d\sigma = 2\pi\chi$, where χ is the \dagger Euler characteristic of the surface. We call $\iint K d\sigma$ the **integral curvature** (or **total Gaussian curvature**).

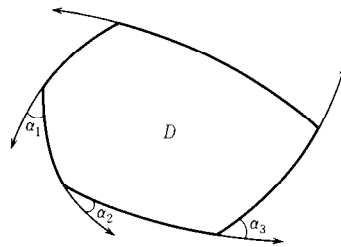


Fig. 7

1. Special Surfaces in E^3

A surface is called a **surface of revolution** if it is generated by a curve C on a plane π when π is rotated around a straight line l in π . Then l and C are called an **axis of rotation** and a **generating curve**, respectively. A surface of revolution having the x_3 -axis as the axis of rotation is given by the equations $x_1 = r \cos \theta$, $x_2 = r \sin \theta$, $x_3 = \varphi(r)$; its first fundamental form is $(1 + \varphi'^2)dr^2 + r^2 d\theta^2$. The section of a surface of revolution by a half-plane through its axis of rotation is called a **meridian**. According as the meridian is a straight line parallel to the axis of rotation or a straight line intersecting the axis nonorthogonally, the surface of revolution is called a **circular cylinder** or a **circular cone**, respectively. If the

meridian is a circle that does not intersect the axis of rotation, it is called a **torus**.

A surface of class C^2 whose mean curvature H vanishes everywhere is called a **minimal surface** (\rightarrow 275 Minimal Submanifolds). A surface of class C^1 realizing a relative minimum of areas among all surfaces of class C^1 with a given closed curve as their boundaries is an \dagger analytic surface such that $H=0$. Conversely, a surface of class C^2 with vanishing mean curvature is an analytic surface. The equation of a surface of revolution with a catenary as its generating line is given by $x_1^2 + x_2^2 = a(e^{x_3/a} + e^{-x_3/a})/2$. This surface is called a **catenoid** and is a minimal surface. Conversely, a minimal surface of revolution is necessarily a catenoid. For a surface obtained by rotating a \dagger Delaunay curve around its base line, the mean curvature H is equal to a constant ($\neq 0$). Conversely, a surface of revolution with nonzero constant mean curvature must be such a surface. A surface with constant Gaussian curvature is called a **surface of constant curvature**, and is a 2-dimensional Riemannian space of constant curvature (\rightarrow 364 Riemannian Manifolds). A non-Euclidean plane can be represented locally as a surface of constant curvature (\rightarrow 285 Non-Euclidean Geometry). Two surfaces of the same constant curvature are locally isometric to each other.

Surfaces of revolution of constant curvature are classified. The simplest surface of constant negative curvature is a **pseudosphere**, which is a surface of revolution obtained by rotating a \dagger tractrix $x_1 = a \cos \varphi$, $x_3 = a \log \tan((\varphi/2) + (\pi/4)) - a \sin \varphi$ ($-\pi/2 < \varphi < \pi/2$) around the x_3 -axis. A surface generated by a 1-parameter family of straight lines is called a **ruled surface**; a hyperboloid of one sheet, a hyperbolic paraboloid, a circular cylinder, and a circular cone are examples. The first two can be regarded as ruled surfaces in two ways. Each of the straight lines that generate a ruled surface is called a **generating line**. A surface consisting of straight lines parallel to a fixed line and passing through each point of a space curve C is called a **cylindrical surface** with the director curve C . A surface generated by a straight line that connects a certain point o with each point of a curve C is called a **conical surface**. Both a cylindrical surface and a conical surface are ruled surfaces such that $K=0$ everywhere. For ruled surfaces we have $K \leq 0$. In particular, a surface such that $H \neq 0$ and $K=0$ everywhere is called a **developable surface**. A developable surface must be either a cylindrical surface, a conical surface, or a tangent surface of a space curve. There exist ruled surfaces that are not developable, for example, hyperboloids of one sheet and hyperbolic paraboloids. A nondevelopable ruled surface is called a **skew surface**. A

ruled surface generated by a straight line that moves under a certain rule intersecting a fixed straight line l orthogonally is called a **right conoid**. If we take l as the x_3 -axis, the surface is given by the equations $x_1 = u \cos v$, $x_2 = u \sin v$, $x_3 = f(v)$. A surface generated by a curve C (C may be chosen as a plane curve) that moves in the direction of a fixed line l with constant velocity and turns around l with certain constant angular velocity is called a **helicoidal surface**. If we take l as the x_3 -axis, the surface is given by the equations $x_1 = u \cos v$, $x_2 = u \sin v$, $x_3 = f(u) + kv$, where k is a constant and $x_3 = f(x_1)$ is the equation of C . In particular, if C is a straight line that intersects l orthogonally, then $f(u)=0$, and the surface is called a **right helicoid** (or **ordinary helicoid**). A right conoid is both a ruled surface and a minimal surface. Conversely, a ruled surface that is also a minimal surface is necessarily a right conoid. A helicoidal surface with a tractrix as the curve C is called a **Dini surface** and is a surface of constant negative curvature. On the normal of a surface S two points q_i ($i=1, 2$) are centers of principal curvature at p . The locus of each of these points is a surface called a **center surface** of S . When S is a sphere, two center surfaces degenerate to a point; if S is a surface of revolution, one of the center surfaces degenerates to the axis of revolution and the other is a certain surface of revolution. If S is general, each of the center surfaces is the locus of an edge of regression of the developable surface generated by normals of S along a line of curvature.

When a 1-parameter family of surfaces S_t is given by the equation $F(x_1, x_2, x_3, t)=0$, a surface E that does not belong to this family is called an **enveloping surface** of the family of surfaces $\{S_t\}$ if E is tangent to some S_t at each point of E , that is, if E and S_t have the same tangent plane. The equation of E is obtained by eliminating t from $F(x_1, x_2, x_3, t)=0$ and $(\partial F/\partial t)(x_1, x_2, x_3, t)=0$. In general, if we denote by $\varphi(x_1, x_2, x_3)=0$ the equation obtained by eliminating t from $F=0$ and $\partial F/\partial t=0$, then the surface defined by $\varphi=0$ is either the enveloping surface of $\{S_t\}$ or the locus of singular points of S_t . The intersection C_{t_0} of the enveloping surface E of $\{S_t\}$ and S_{t_0} is a curve defined by $F(x_1, x_2, x_3, t_0)=0$, $(\partial F/\partial t)(x_1, x_2, x_3, t_0)=0$. We call C_{t_0} a **characteristic curve** of $\{S_t\}$. Since $\{C_t\}$ is a family of curves on the enveloping surface E , there may exist an envelope F on E . In such a case, F is called the **line of regression** of $\{S_t\}$. The equation of F is obtained by eliminating t from $F=0$, $\partial F/\partial t=0$, and $\partial^2 F/\partial t^2=0$. In particular, the enveloping surface of a family of planes is a developable surface, and their characteristic curves are straight lines. Moreover, the line of

regression coincides with the line of regression of the tangent surface.

If there exists a diffeomorphism between two surfaces such that first fundamental forms at each pair of corresponding points are proportional, then the surfaces are said to be in a **conformal correspondence**. In particular, when the proportionality factor is a constant, they are said to be in a **similar** (or **homothetic**) **correspondence**. There exists a local conformal correspondence between any analytic surface and a plane. Namely, if we choose suitable parameters, we can reduce the first fundamental form of any analytic surface to the form $A(\xi, \eta)(d\xi^2 + d\eta^2)$. Such parameters are called **isothermal parameters**. From the existence of isothermal parameters we can see that there exists a local conformal correspondence between any two analytic surfaces. The assumption of analyticity in these theorems is not necessary [10]. If there exists a diffeomorphism between two surfaces under which geodesics are mapped to geodesics, then the surfaces are said to be in **geodesic correspondence**. A surface has a locally geodesic correspondence with a plane if and only if it is a surface of constant curvature. If two surfaces are in geodesic correspondence, then with respect to parameters with the same values at corresponding points, we have the relation

$$\overline{\left\{ \begin{matrix} \alpha \\ \beta_\gamma \end{matrix} \right\}} = \left\{ \begin{matrix} \alpha \\ \beta_\gamma \end{matrix} \right\} + \delta_\beta^\alpha A_\gamma + \delta_\gamma^\alpha A_\beta$$

for coefficients of connections of the two surfaces.

By the [†]Alexander-Pontryagin duality theorem, a submanifold M in E^3 that is homeomorphic to S^2 divides E^3 into two domains, and two points belonging to different domains cannot be connected by a broken segment unless the segment meets the surface. Such a manifold M is called a **closed surface**. One of the two domains consists of those points with bounded distance from a point belonging to the domain. Such a domain is called the interior of the closed surface M . If the set M^* consisting of M and its interior is convex in E^3 , the surface M is called a **closed convex surface** (or **ovaloid**).

The Gaussian curvature of an ovaloid cannot be negative at any point. A closed surface with $K > 0$ must be an ovaloid. Moreover, it is known that on any closed surface there exists at least one point where $K > 0$ (J. Hadamard). If there exists no umbilical point and K is strictly positive in a domain on a surface, then the two principal curvatures regarded as continuous functions on the domain cannot take their local maximum and local minimum values at the same point (D. Hilbert). A com-

pact, connected surface of class C^4 with constant Gaussian curvature is a sphere. A closed surface with $K > 0$ and $H = \text{constant}$ is a sphere (H. Liebmann).

A problem proposed by H. Hopf asks whether an orientable compact surface with constant mean curvature is a sphere. In connection with this problem, Hopf showed that a closed orientable surface of class C^3 of genus zero with constant mean curvature is a sphere. If there is a certain relation $W(k_1, k_2) = 0$ between the two principal curvatures k_1, k_2 ($k_1 \geq k_2$) of a surface, the surface is called a **Weingarten surface** (or **W-surface**). There are many interesting results for W -surfaces. As an extension of the convex surface, tight immersions have been studied (\rightarrow 365 Riemannian Submanifolds).

The Gaussian curvature K is invariant under isometries. Hence a sphere is transformed to a sphere by each isometry. This fact is sometimes described as the **rigidity** of a sphere. More generally, if two ovaloids are isometric, then they are congruent (**Cohn-Vossen's theorem**). It is known that if we remove a small circular disk from a sphere, then the remaining portion of the sphere is isometrically deformable. On the existence of closed geodesics on ovaloids, G. D. Birkhoff proved the following theorem: There exist at least three closed geodesics on any ovaloid of class C^3 . It is also known that there exist surfaces of revolution that are not spheres but whose geodesics are all closed (\rightarrow 178 Geodesics). On a hyperbolic non-Euclidean compact [†]space form of genus p ($p \geq 2$) there exists a geodesic whose points are everywhere dense in it (E. Hopf) (for the ergodicity of flows along geodesics on this surface \rightarrow 136 Ergodic Theory; also 126 Dynamical Systems).

J. Singular Points of a Surface

Suppose that a neighborhood of a point p_0 of a surface S in E^3 is given by a certain vector-valued function f of class C^r as $r = \tilde{r}(u, v)$. Then a point p_0 where two vectors $(\partial \tilde{f} / \partial u)_{p_0}, (\partial \tilde{f} / \partial v)_{p_0}$ are linearly independent is called a **regular point**. A point on S that is not regular is called a **singular point**. If for suitable parameters we have $(\partial \tilde{f} / \partial u)_{p_0} = 0$ but $(\partial \tilde{f} / \partial v)_{p_0}, (\partial^2 \tilde{f} / \partial u^2)_{p_0}, (\partial^2 \tilde{f} / \partial u \partial v)_{p_0}$ are linearly independent, then such a singular point is called a **semiregular point**. In general, shapes of neighborhoods of singular points are extremely complicated. However, we note the following: (i) by a small deformation of the function f (and its derivatives of orders at most r) we can reduce p_0 to a regular or semiregular point of the deformed surface; (ii) if p_0 is semiregular, we can choose

suitable parameters and curvilinear coordinates of class C^r in E^3 near p_0 so that the surface S in the neighborhood of the origin p_0 is expressed by the equations $x_1 = u^2$, $x_2 = v$, $x_3 = uv$ (H. Whitney's theorem [15]). (The higher-dimensional case has also been considered (Whitney [16]).)

References

- [1] W. Blaschke, *Vorlesungen über Differentialgeometrie I*, Springer, 1924 (Chelsea, 1967).
- [2] A. Duschek and W. Mayer, *Lehrbuch der Differentialgeometrie I*, Teubner, 1930.
- [3] L. P. Eisenhart, *An introduction to differential geometry*, Princeton Univ. Press, 1940; revised edition, 1947.
- [4] W. Klingenberg, *Eine vorlesung über differentialgeometrie*, Springer, 1973.
- [5] W. Fenchel, On the differential geometry of closed space curves, *Bull. Amer. Math. Soc.*, 57 (1951), 44–54.
- [6] S. S. Chern, *Curves and surfaces in Euclidean space*, Studies in global geometry and analysis, Studies in math. 4, Math. Assoc. Amer.
- [7] I. Fary, Sur la courbure totale d'une courbe gauche faisant un noeud, *Bull. Soc. Math. France*, 77 (1949), 128–138.
- [8] J. W. Milnor, On the total curvature of knots, *Ann. Math.*, (2) 52 (1950), 248–257.
- [9] E. Cartan, La théorie des groupes finis et continus et la géométrie différentielle traitées par la méthode du repère mobile, Gauthier-Villars, 1937.
- [10] L. Bers, Riemann surfaces, *Courant Inst. Math. Sci.*, 1957–1958.
- [11] S. Sternberg, *Lectures on differential geometry*, Prentice-Hall, 1964.
- [12] A. D. Aleksandrov, *Intrinsic geometry of convex surfaces* (in Russian), Gostekhizdat, 1948; German translation, *Die innere Geometrie der konvexen Flächen*, Akademie-Verlag, 1955.
- [13] S. S. Chern, *Topics in differential geometry*, Lecture notes, Institute for Advanced Study, Princeton, 1951.
- [14] H. Hopf, *Zur Differential Geometrie geschlossener Flächen in Euklidischen Raum*, Convegno Internazionale di Geometria Differenziale, Italy, 1953, 45–54 (Cremona, 1954).
- [15] H. Whitney, The singularities of a smooth n -manifold in $(2n-1)$ -space, *Ann. Math.*, (2) 45 (1944), 247–293.
- [16] H. Whitney, Singularities of mappings of Euclidean spaces, *Symposium Internacional de Topologia Algebraica*, Universidad Nacional Autonoma de Mexico and UNESCO, 1958, 285–301.

- [17] M. P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, 1976.
- [18] T. J. Willmore, *An introduction to differential geometry*, Clarendon Press, 1959.
- [19] N. J. Hicks, *Notes on differential geometry*, Van Nostrand, 1964.
- [20] D. Laugwitz, *Differential and Riemannian geometry*, Academic Press, 1965.
- [21] B. O'Neill, *Elementary differential geometry*, Academic Press, 1966.
- [22] J. J. Stoker, *Differential geometry*, Wiley, 1969.

112 (XII.15) Differential Operators

A. Definition

A mapping (or an operator) A of a function space F_1 to a function space F_2 is said to be a **differential operator** if the value $f(x)$ of the image $f = Au$ ($u \in F_1, f \in F_2$) at each point x is determined by the values at x of u and a finite number of its derivatives. If u and f are distributions, the definition applies with the derivative interpreted in the sense of distributions (\rightarrow 125 Distributions and Hyperfunctions). In this article we restrict ourselves to the case of linear differential operators and consider only those of the form

$$P(x, D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha, \quad (1)$$

where α denotes n -tuples $(\alpha_1, \alpha_2, \dots, \alpha_n)$ of nonnegative integers, called **multi-indices**; $|\alpha|$ the **length** of α : $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$; and D^α the differential operator $D^\alpha = D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n}$, with $D_j = (-i)\partial/\partial x_j$. The coefficient $(-i)$ is sometimes omitted. The coefficients $a_\alpha(x)$ are functions defined on an open set Ω in n -dimensional space. We call $P(x, D)$ an **ordinary differential operator** if the dimension n of Ω is 1 and a **partial differential operator** if $n \geq 2$. Ordinary differential operators and partial differential operators behave quite differently in many respects.

We set

$$P(x, \xi) = \sum_{|\alpha| \leq m} a_\alpha(x) \xi^\alpha, \quad \xi^\alpha = \xi_1^{\alpha_1} \dots \xi_n^{\alpha_n},$$

where $\xi = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbf{R}^n$ or \mathbf{C}^n . The **order** of $P(x, D)$ is the greatest integer $|\alpha|$ for which $a_\alpha(x) \neq 0$. In expression (1) m is assumed to be equal to the order, and in that case

$$P_m(x, D) = \sum_{|\alpha|=m} a_\alpha(x) D^\alpha$$

is called the **principal part** of $P(x, D)$, and the corresponding polynomial $P_m(x, \xi)$ the **characteristic polynomial**.

Differential operators have been investigated for a long time in connection with the linear differential equations

$$P(x, D)u(x) = f(x), \quad x \in \Omega. \quad (2)$$

Except for ordinary differential operators, however, it is rather recently that the properties of such operators have been studied from the general viewpoint.

We denote a differential operator with constant coefficients by $P(D)$. In general, $P(x, D)$ is assumed to be a linear differential operator with coefficients that are C^∞ -functions. However, many of the results for C^∞ -coefficients also hold when the coefficients are sufficiently differentiable. (For \dagger function spaces $\mathcal{D}(\Omega)$, $\mathcal{D}'(\Omega)$, $\mathcal{E}(\Omega)$, $\mathcal{E}'(\Omega)$, $C(\Omega)$, $C^\infty(\Omega)$, $C_0^\infty(\Omega)$, $L_p(\Omega)$, $\mathcal{A}(\Omega)$, $\mathcal{B}(\Omega)$, etc., \rightarrow 125 Distributions and Hyperfunctions, 168 Function Spaces).

Differential operators are classified according to their properties. The most important are the elliptic, hyperbolic, and parabolic types. A differential operator $P(x, D)$ is said to be an **elliptic** operator if the characteristic polynomial $P_m(x, \xi)$ has no real zero except for $\xi = 0$ for each $x \in \Omega$. Typical examples are the Laplacian $\Delta = -(D_1^2 + \dots + D_n^2)$ and the Cauchy-Riemann operator $\partial/\partial\bar{z} = (1/2)(\partial/\partial x + i\partial/\partial y)$.

A differential operator is said to be **hyperbolic** if the associated Cauchy problem is well posed (\rightarrow 325 Partial Differential Equations of Hyperbolic Type). The d'Alembertian $D_1^2 - (D_2^2 + \dots + D_n^2)$ is an example. A differential operator of the form $\partial/\partial t + P(t, x, D_x)$ is called **parabolic** if $P(t, x, D_x)$ is strongly elliptic (\rightarrow Section G) in x . The heat operator $iD_{n+1} - \Delta$ is typical.

These three types of operators appear most often in applications, and if $n = 2$, then any operator of order 2 with real coefficients in $(\partial/\partial x_j)^2$ belongs to one of them at a generic point. In other cases, however, there are differential operators that do not belong to any of them.

B. Fundamental Solutions

If a differential operator $P(x, D)$ with $\mathcal{D}(\Omega)$ as its domain has a left inverse F that is expressed as an \dagger integral operator with \dagger kernel distribution (in $\mathcal{D}'_{x,y}$; \rightarrow 125 Distributions and Hyperfunctions F), then the kernel is said to be a **fundamental solution** (or **elementary solution**). F is usually a right inverse of the weak extension (\rightarrow Section F) of $P(x, D)$ and maps $\mathcal{D}(\Omega)$ into $\mathcal{E}(\Omega)$. The image is mapped to the original function by $P(x, D)$. Nevertheless, F is not a genuine right inverse, and hence the fundamental solution is not unique if it exists.

When $P(D)$ is a differential operator with constant coefficients, we call a distribution $E(x)$ a **fundamental solution** if it satisfies

$$P(D)E(x) = \delta(x), \quad (3)$$

where $\delta(x)$ is \dagger Dirac's distribution (δ function). If $E(x)$ is a fundamental solution in this sense, then $F(x, y) = E(x - y)$ is the kernel of a left inverse of $P(D)$ and is a fundamental solution in the sense of the preceding paragraph.

Every differential operator $P(D)$ with constant coefficients has a fundamental solution in the sense of (3) (**Ehrenpreis-Malgrange theorem**; see L. Hörmander [4] for a proof).

General operators $P(x, D)$ with variable coefficients do not necessarily have fundamental solutions. However, if $P(x, D)$ belongs to one of the classical types of operators (elliptic, hyperbolic, or parabolic), then it has a fundamental solution at least locally. (See F. John [8] for elliptic operators, J. Leray [11] for strongly hyperbolic operators, and S. Mizohata [12] and S. D. Eidel'man [9] for parabolic operators.) Leray has generalized John's method to strongly hyperbolic operators in an enormous work [13].

C. Ranges of Differential Operators

Let $P(D)$ be a differential operator with constant coefficients. Then it follows from the Ehrenpreis-Malgrange theorem that $P(D)\mathcal{D}'(\Omega) \supset \mathcal{D}(\Omega)$ holds for any open set Ω . However, there are differential operators $P(x, D)$ with variable coefficients such that for any Ω , $P(x, D)\mathcal{D}'(\Omega) \not\supset \mathcal{D}(\Omega)$. H. Lewy first devised such an example:

$$P(x, D) = -iD_1 + D_2 - 2(x_1 + ix_2)D_3.$$

Let $C_{2m-1}(x, D)$ be the homogeneous part of order $2m-1$ of the commutator

$$P(x, D)\overline{P(x, D)} - \overline{P(x, D)}P(x, D).$$

Then in order that $P(x, D)\mathcal{D}'(\Omega) \supset \mathcal{D}(\Omega)$, it is necessary that

$$P_m(x, \xi) = 0 \quad \text{imply} \quad C_{2m-1}(x, \xi) = 0$$

for all $x \in \Omega$, $\xi \in \mathbb{R}^n$ (**Hörmander's theorem** [4]).

When $P(x, D)$ is a differential operator that does not satisfy this condition (e.g., Lewy's operator), choose an $f(x) \in \mathcal{D}(\Omega)$ that is not in $P(x, D)\mathcal{D}'(\Omega)$. Then the differential equation (2) has no distribution solutions at all. P. Schapira extended this result to the case of \dagger hyperfunctions (also \rightarrow 274 Microlocal Analysis).

Concerning the ranges of differential operators $P(D)$ with constant coefficients, we have the following detailed results due to L. Ehrenpreis [28], B. Malgrange [14], and Hörmander [4].

An open set Ω is said to be ***P*-convex** for a differential operator $P(D)$ if for each compact set $K \subset \Omega$ there exists a compact set $K' \subset \Omega$ such that $\varphi \in C_0^\infty(\Omega)$ and $\text{supp } P(-D)\varphi \subset K$ imply $\text{supp } \varphi \subset K'$. Convex sets are *P*-convex for any $P(D)$. All open sets are *P*-convex if and only if $P(D)$ is an elliptic operator.

Theorem: The following conditions are equivalent: (i) Ω is *P*-convex; (ii) $P(D)\mathcal{D}'(\Omega) \supset \mathcal{E}(\Omega)$; (iii) $P(D)\mathcal{E}(\Omega) = \mathcal{E}(\Omega)$. Property (iii), the Mittag-Leffler theorem, and the solvability of †Cousin's first problem for the solutions of $P(D)u = 0$ are equivalent.

An open set Ω is said to be **strongly *P*-convex** if for each compact set $K \subset \Omega$ there exists a compact set K' such that $\mu \in \mathcal{E}'(\Omega)$ and $\text{supp } P(-D)\mu \subset K$ imply $\text{supp } \mu \subset K'$; and $\mu \in \mathcal{E}'(\Omega)$ and $\text{sing supp } P(-D)\mu \subset K$ imply $\text{sing supp } \mu \subset K'$; where the **singular support** of μ is the closure of the set of all points at which μ is not a C^∞ -function. Convex sets are strongly *P*-convex, and strongly *P*-convex sets are *P*-convex.

Theorem: Ω is strongly *P*-convex if and only if $P(D)\mathcal{D}'(\Omega) = \mathcal{D}'(\Omega)$.

R. Harvey has shown that every domain Ω is *P*-convex in the sense of hyperfunctions, i.e., the equation $P(D)u = f$ always has a hyperfunction solution u on Ω for any hyperfunction f on Ω . For the †real analytic functions $\mathcal{A}(\Omega)$, however, $P(D)\mathcal{A}(\Omega) = \mathcal{A}(\Omega)$ does not hold for convex open set Ω in general. Hörmander [15] gave a necessary and sufficient condition for $P(D)$ and Ω in order that this hold.

D. Hypoellipticity

A differential operator $P(x, D)$ is called **hypoelliptic** in Ω if for any distribution $u(x) \in \mathcal{D}'(\Omega)$, $Pu \in C^\infty(\Omega)$ implies $u \in C^\infty(\Omega)$. Further, a differential operator with real analytic coefficients $P(x, D)$ is called **analytically hypoelliptic** in Ω if P is hypoelliptic and if for any distribution $u(x) \in \mathcal{D}'(\Omega)$, $Pu \in \mathcal{A}(\Omega)$ implies $u(x) \in \mathcal{A}(\Omega)$.

There are two fundamental facts about such operators. Let P have constant coefficients; then $P(D)$ is hypoelliptic if and only if $P(\xi + i\eta) = 0$ and $|\xi + i\eta| \rightarrow \infty$ imply that $|\eta| \rightarrow \infty$ (**Hörmander's theorem** [4]). Furthermore, $P(D)$ is analytic hypoelliptic if and only if P is elliptic (**Petrovskii's theorem** [16]). The heat operator is not elliptic, but hypoelliptic. If $P(D)$ is elliptic, then actually any hyperfunction $u(x) \in \mathcal{B}(\Omega)$ such that $Pu \in \mathcal{A}(\Omega)$ is real analytic (R. Harvey, G. Bengel). On the other hand, if $P(D)$ is not elliptic, there is a hyperfunction solution u of $Pu = 0$ that is not a distribution.

Strictly speaking, the notion of the hypoellipticity for general differential operators was first formulated explicitly by L. Schwartz

[27]. Before that time, D. Hilbert, E. E. Levi and K. O. Friedrichs, and others investigated this problem for some elliptic operators, and the hypoellipticity was called **Weyl's lemma** (\rightarrow 323 Partial Differential Equations of Elliptic Type).

Similarly to the constant coefficient case the following two theorems are fundamental: Elliptic and parabolic operators $P(x, D)$ with C^∞ coefficients are hypoelliptic (Schwartz [27], Mizohata [12]). Elliptic operators $P(x, D)$ with real analytic coefficients are analytic hypoelliptic (I. G. Petrovskii [16], C. B. Morrey and L. Nirenberg). The latter holds also for hyperfunction solutions (M. Sato and Schapira). Moreover, the following result is known: If for each compact set K in an open set Ω , there exists a constant C such that $\|P^k u\|_K \leq C^{k+1}(mk)!$, then $u \in \mathcal{A}(\Omega)$, where $\|\cdot\|_K$ denotes the L_p -norm on K (H. Komatsu, T. Kotake, and M. S. Narasimhan).

However, as seen by the example $D_1 + ix_1^k D_2$ in \mathbf{R}^2 (k is even), the analytic hypoellipticity also holds for nonelliptic operators. Such an operator is called a subelliptic operator; subelliptic operators have been investigated by Hörmander, Yu. V. Egorov, F. Trèves, and others [19].

Hörmander has obtained a fairly complete result on the hypoellipticity of the operators of the form

$$L = \sum_{i=1}^r X_i^2 + X_0 + c(x),$$

where X_0, \dots, X_r are homogeneous first-order differential operators with real coefficients ([17]; O. A. Oleinik and E. V. Radkevich [18]). Hypoellipticity was investigated extensively after the introduction of pseudodifferential operators and Fourier integral operators (\rightarrow 345 Pseudodifferential Operators).

E. Differential Operators in Banach Spaces

We consider differential operators $P(x, D)$ defined on a domain Ω as operators in the function spaces $C(\Omega)$ or $L_p(\Omega)$. Differential operators of order $m \geq 1$ are always unbounded operators in the Banach space $X = C(\Omega)$ or $L_p(\Omega)$. Moreover, their domains of definition as operators in X are not generally determined uniquely by the expressions $P(x, D)$ as differential operators.

$P(x, D)$ is a linear operator that maps $C_0^\infty(\Omega)$ into X . This operator has a closed extension. The minimal closed extension P_0 is called the **minimal operator** of $P(x, D)$ in X . We have $u \in \mathcal{D}(P_0)$ and $P_0 u = f$ if and only if there exists a sequence $\varphi_n \in C_0^\infty(\Omega)$ such that $\varphi_n \rightarrow u$, $P(x, D)\varphi_n \rightarrow f$. On the other hand, the

closed linear operator P_1 whose domain is the set of all $u \in X$ such that $P(x, D)u \in X$ in the sense of distribution is called the **maximal operator** (or **weak extension**) of $P(x, D)$. We have $u \in \mathcal{D}(P_1)$ and $P_1 u = f$ if and only if $\langle u, {}^t P(x, D)\varphi \rangle = \langle f, \varphi \rangle$ for any $\varphi \in C_0^\infty(\Omega)$, where ${}^t P(x, D)$ is the **transposed operator**

$${}^t P(x, D)\varphi(x) = \sum_{|z| \leq m} (-D)^z (a_z(x)\varphi(x)).$$

Integration by parts shows that P_1 is an extension of P_0 and that when X is the t dual space of a space Y , the weak extension P_1 in X is the t dual of the minimal operator of ${}^t P(x, D)$ in Y . Let $X = L_p(\Omega)$ ($1 < p < \infty$), Ω be a bounded open set with smooth boundary, and $P(D)$ have constant coefficients. Then P_1 coincides with the smallest closed extension of the operator $P(D)$ having as its domain the set of all $u \in C^\infty(\Omega) \cap X$ such that $P(D)u \in X$. The latter closed extension is called the **strong extension**. The difference between the weak and the strong extension is not obvious in the variable coefficient case.

P_0 coincides with P_1 when Ω is the entire space and $P(x, D)$ is an elliptic operator whose coefficients are constants or close to constants (J. Peetre, *Medd. Lunds Univ. Mat. Sem.*, 16 (1959); T. Ikebe and T. Kato, *Arch. Rational Mech. Anal.*, 9 (1962)). In general, we have $P_0 \neq P_1$. Let $P(x, D)$ be an ordinary differential operator with bounded coefficients such that $|a_m(x)| \geq \delta > 0$ and Ω be the bounded interval (a, b) . Then the domain of P_1 coincides with the set of all $(m-1)$ -times continuously differentiable functions u such that the $(m-1)$ st derivative is absolutely continuous and $P(x, D)u \in X$, while the domain of P_0 is the set of all functions u which satisfy in addition the boundary conditions

$$\begin{aligned} u(a) &= u'(a) = \dots = u^{(m-1)}(a) \\ &= u(b) = \dots = u^{(m-1)}(b) = 0. \end{aligned}$$

(Moreover, $u^{(m)}(a) = u^{(m)}(b) = 0$ when $X = C(a, b)$.)

Let $G(P_0), G(P_1) (\subset X \times X)$ be the t graphs of P_0, P_1 . Then the quotient space $\mathcal{B} = G(P_1)/G(P_0)$ is called the **boundary space**, and an element of the dual \mathcal{B}' of \mathcal{B} , i.e., a continuous linear functional on $G(P_1)$ which vanishes on $G(P_0)$, is called a **boundary value** relative to $P(x, D)$. For the ordinary differential operators discussed above, the boundary space is the set of all linear combinations of $u^{(i)}(a)$ and $u^{(j)}(b)$. When $P(x, D)$ is an ordinary differential operator, we can explicitly determine the boundary values also in the case where the interval (a, b) is infinite, the coefficients $a_z(x)$ are not bounded, or $a_m(x) \rightarrow 0$ ($x \rightarrow a, b$); and we can show that \mathcal{B} is finite-dimensional. When $P(x, D)$ is a partial differential operator, \mathcal{B} is

generally of infinite dimension, and the concrete forms of the elements of \mathcal{B} and \mathcal{B}' are not known. However, we have some information by M. I. Vishik (*Amer. Math. Soc. Transl.*, (2) 24 (1963)) about the boundary values of elliptic operators of the second order. Combining this with the results by J.-L. Lions and E. Magenes (*J. Anal. Math.*, 11 (1963)), we can obtain information for elliptic operators of higher order.

F. Differential Operators with Boundary Conditions

A closed operator between the minimal operator P_0 and the maximal operator P_1 is determined by designating a closed subspace B of the boundary space \mathcal{B} . This operator is called the **operator with the boundary condition B** . Particularly important are boundary conditions expressed in the form

$$Q_i(x, D)u(x) = 0, \quad x \in \partial\Omega, \quad i = 1, \dots, k, \quad (4)$$

with differential operators $Q_i(x, D)$ ($i = 1, \dots, k$) defined on the boundary $\partial\Omega$ of Ω .

When $P(x, D)$ is an ordinary differential operator defined on a finite interval and the orders of Q_i are at most $m-1$ (or m), (4) always has a definite meaning. However, for partial differential operators, we need an interpretation of (4), i.e., (4) does not necessarily determine the subspace B of \mathcal{B} uniquely.

Let P_s be the smallest closed extension of $P(x, D)$ with $\{u \in C^\infty(\Omega) \cap X \mid Q_i(x, D)u(x) = 0, x \in \partial\Omega; P(x, D)u \in X\}$ as its domain. P_s is called the **strong extension** of the differential operator $P(x, D)$ with boundary condition (4).

On the other hand, when Ω, P , and Q_i satisfy suitable conditions, we can define the transposed differential operator ${}^t P(x, D)$ with the transposed boundary operators $R_j(x, D)$ ($j = 1, \dots, l$). Namely, there are differential operators $R_j(x, D)$ on the boundary such that a necessary and sufficient condition for $u \in C^\infty(\bar{\Omega})$ to satisfy (4) and $Pu(x) = f(x)$ is

$$\int_{\Omega} f(x)v(x)dx = \int_{\Omega} u(x){}^t P v(x)dx \quad (5)$$

for all $v(x) \in C^\infty(\bar{\Omega})$ with the boundary conditions $R_j v = 0, x \in \partial\Omega$. Then the operator P_w defined by $P_w u(x) = f(x)$ for the pairs $u(x), f(x) \in X$ satisfying (5) is called the **weak extension** of the differential operator $P(x, D)$ with boundary condition (4). As in the case of operators without boundary conditions, the weak extension is an extension of the strong extension, and generally is the dual of the strong extension in the dual space X' of the transposed differential operator with the transposed boundary condition.

Regularity up to the boundary. The fundamental problems for the differential operator $P(x, D)$ with the homogeneous boundary condition (4) are to determine, for both the strong and the weak extensions, the spaces of solutions of the homogeneous equations $Pu=0$ and their ranges. The problems mostly reduce to determining when the strong and the weak extensions coincide and, including this, also to the problem of regularity on the closed domain $\bar{\Omega}$ containing the boundary of the solutions u of the equation $P_w u = f$.

This problem was solved by Nirenberg for strongly elliptic operators in $L^2(\Omega)$ with the Dirichlet boundary condition

$$\partial^{j-1} u(x) / \partial n^{j-1} = 0, \quad j = 1, 2, \dots, m/2,$$

and generalized later by F. Browder, M. Schechter [20], S. Agmon, Lions, and others for elliptic operators in $L_p(\Omega)$ with a kind of coercive boundary condition (\rightarrow Section H).

Consequently, when Ω is bounded and smooth, P_w is equal to P_s for those operators. Write P for P_w . Then the space $N(P)$ of the solutions of $Pu=0$ is a subspace of finite dimension, and the range $R(P)$ is a closed subspace of finite codimension. In particular, it follows that the index of P , $\dim N(P) - \text{codim } R(P)$, is finite.

G. Strongly Elliptic Operators

A differential operator $P(x, D)$ is said to be **strongly elliptic** if its characteristic polynomial satisfies

$$\text{Re } P_m(x, \xi) \geq C |\xi|^m > 0, \quad \xi \neq 0.$$

Many of the elliptic operators, such as the Laplacian, that appear in applications are strongly elliptic. L. Gårding treated the boundary value problem with the Dirichlet condition (in the generalized sense) for strongly elliptic operators. His work initiated the general study of differential operators (\rightarrow 323 Partial Differential Equations of Elliptic Type). His theory is based on the following inequality, called **Gårding's inequality**:

$$\|\nabla^{m/2} u\|_{L_2}^2 \leq C \left(\text{Re} \int Pu \cdot \bar{u} dx + \|u\|_{L_2}^2 \right),$$

$$u \in C_0^\infty(\Omega).$$

H. Coercive Boundary Conditions

The boundary condition (4) is said to be **coercive** if

$$\|\nabla^m u\| \leq C(\|Pu\| + \|u\|) \quad (6)$$

holds for any $u \in C^\infty(\bar{\Omega})$ that satisfies (4). In

order that a differential operator P have a coercive boundary condition, it is necessary that it be a special type of elliptic operator. In this case, Agmon, N. Aronszajn, Schechter, and others found conditions under which (4) is coercive. Agmon, A. Douglis, and Nirenberg show that the inequality (6) holds in $L_p(\Omega)$ and in the normed spaces of Hölder continuous functions under a suitable condition. The classical boundary conditions $au + b\partial u/\partial n = 0$ ($a \geq 0, b \geq 0, a + b = 1$) are coercive for elliptic operators of the second order (\rightarrow 323 Partial Differential Equations of Elliptic Type). However, problems remain when the coefficients of $Q_j(x, D)$ are discontinuous. In order to have coincidence of the strong and weak extensions or regularity up to the boundary, it is necessary neither that P be elliptic nor that the boundary condition be coercive. But it is not known to what extent these conditions can be weakened. At present, major contributions are Hörmander's work (*Acta Math.*, 99 (1958)) dealing with operators with constant coefficients and flat boundaries, and works by J. J. Kohn [21], Nirenberg, and Hörmander concerning noncoercive boundary conditions. The latter works are connected with the theory of several complex variables, and have attracted much attention.

I. Self-Adjoint Extension

One of the fundamental problems in the case $X = L_2(\Omega)$ is whether the minimal operator P_0 has a self-adjoint extension. P_0 is symmetric if and only if $P(x, D)$ is **formally self-adjoint**: $P(x, D) = {}^t P(x, D)$. Under this condition the boundary space \mathcal{B} turns out to be the direct sum of two subspaces $\mathcal{B}_\pm = \{(x, P_1 x) \mid x \in \mathcal{D}(P_1), P_1 x = \pm ix\} + G(P_0)$. The numbers $n_\pm = \dim \mathcal{B}_\pm$ are called the **deficiency indices** of P_0 , and P_0 has a self-adjoint extension if and only if $n_+ = n_-$.

H. Weyl gave a method for computing n_\pm for the **Sturm-Liouville operators**:

$$P(x, D) = - \left(\frac{d}{dx} p(x) \frac{d}{dx} \right) + q(x), \quad x \in (a, b).$$

We say that a (resp. b) is of **limit circle type** if the solutions $u(x)$ of $P(x, D)u(x) + lu(x) = 0$ ($l \in \mathbb{C}$) always belong to L_2 in a neighborhood of a (resp. b) and of **limit point type** if a solution does not belong to L_2 . This classification does not depend on the choice of $l \in \mathbb{C}$. (1) If both a and b are of limit point type, then $n_+ = n_- = 0$ and hence P_0 is self-adjoint. (2) If a is of limit circle type and b is of limit point type, then $n_+ = n_- = 1$, and the self-adjoint extensions of P_0 are the operators P_α that are obtained from P_1 by assigning the boundary

condition

$$p(a)u'(a)\cos\alpha + u(a)\sin\alpha = 0.$$

The same is true when a and b are interchanged. (3) If both a and b are of limit circle type, then $n_+ = n_- = 2$, and we can impose two boundary conditions to obtain the self-adjoint extensions.

These results have been extended by K. Kodaira [23] and N. Dunford and J. Schwartz [3] to the case of ordinary differential operators of order m . There are formally self-adjoint operators that have no self-adjoint extensions. For example, the operator $-id/dx$ in $L_2(0, \infty)$ has deficiency indices $n_+ = 1 \neq n_- = 0$.

For partial differential operators, it is difficult to determine explicitly all self-adjoint extensions of a given formally self-adjoint operator because the boundary space is complicated. If the boundary condition (4) is formally self-adjoint and coercive, then it follows from the results of Schechter and others that the differential operator with boundary condition (4) is self-adjoint. Furthermore, conditions under which P_0 is self-adjoint or has a self-adjoint extension are known. The following theorem is often used as a condition of the latter type. If a \dagger symmetric operator defined on a dense subspace of a Hilbert space X is positive definite:

$$(Tx, x) \geq 0, \quad x \in \mathcal{D}(T),$$

then there is a positive definite self-adjoint extension \tilde{T} (Friedrichs's theorem). The self-adjoint extension obtained by this theorem is called the **Friedrichs extension**.

J. Generators of Semigroups

From the point of view of probability theory, W. Feller investigated the extensions of the Laplacian d^2/dx^2 and similar operators that are the generators of order-preserving semigroups. Recently various attempts have been made to generalize his results to the multi-dimensional case (\rightarrow 115 Diffusion Processes; 378 Semigroups of Operators and Evolution Equations).

P. D. Lax and A. N. Milgram proved that if $P(x, D)$ is a strongly elliptic operator, then $-P(x, D)$ with the Dirichlet condition in $L_2(\Omega)$ is the generator of a semigroup [1].

K. Boundary Value Problems

There are two methods of solving the inhomogeneous boundary value problem

$$P(x, D)u(x) = f(x), \quad x \in \Omega,$$

$$Q_i(x, D)u(x) = g_i(x), \quad x \in \partial\Omega, \quad i = 1, \dots, k.$$

In the first method, we take a function $v(x)$ that satisfies $Q_i(x, D)v(x) = g_i(x)$ and reduce the problem to the homogeneous one for $u_0 = u - v$. In the second method, we consider the pair $P(x, D)$ and $Q_i(x, D)$ as an operator that maps a function u to the pair of functions $(Pu, Q_i u)$ and investigate it directly. The latter method was adopted by Peetre and Hörmander [4].

L. Estimates in Weighted Spaces

J. F. Treves, Hörmander [4], and H. Kumano-go obtained estimates similar to (6) in L_p -spaces relative to the weighted measure $w_i(x)dx$ instead of the usual L_p -spaces, and applied them to the proof of the uniqueness of Cauchy problems for differential equations with variable coefficients (\rightarrow 321 Partial Differential Equations (Initial Value Problems)). Hörmander [22] applied similar estimates to the proof of the \dagger fundamental theorems of Stein manifolds.

M. Eigenfunction Expansions

When a self-adjoint operator P in the Hilbert space $L_2(\Omega)$ is a self-adjoint extension of a differential operator $P(x, D)$, the \dagger spectral decomposition of P is concretely expressed by the expansion of functions $u \in L_2(\Omega)$ into eigenfunctions of $P(x, D)$.

If Ω is bounded, $P(x, D)$ is an elliptic operator defined on a neighborhood of $\bar{\Omega}$, and the boundary condition is coercive, then the \dagger spectrum of P is composed solely of eigenvalues, and the eigenvectors of P are eigenfunctions of $P(x, D)$ in the classical sense and are of class C^∞ up to the boundary.

N. Asymptotic Distribution of Eigenvalues

If $P = -\Delta$, the number $v(\lambda)$ of eigenvalues less than λ satisfies the asymptotic relation

$$v(\lambda) \sim \frac{\lambda^{n/2} A}{2^{n-1} \pi^{n/2} n \Gamma(n/2)}, \quad \text{as } \lambda \rightarrow \infty$$

regardless of the shape of the domain and the boundary condition, where n is the dimension and A is the volume of Ω [2]. This was first proved by Weyl and extended by T. Carleman, Gårding, and others to the case of operators of higher order with variable coefficients (\rightarrow 323 Partial Differential Equations of Elliptic Type).

O. Weyl-Stone-Titchmarsh-Kodaira Theory

When Ω is unbounded or the coefficients of $P(x, D)$ have singularities near the boundary of Ω , the spectrum of P may have a continuous part.

Let $P(x, D)$ be an ordinary differential operator on an interval (a, b) . Then for each $\lambda \in \mathbb{C}$ the equation $(P(x, D) - \lambda)\varphi(x) = 0$ has m linearly independent solutions $\varphi_k(x, \lambda)$ ($k = 1, \dots, m$), and any solution is represented as a linear combination of them. Weyl and M. H. Stone obtained the spectral decomposition of P (of the second order) in the form

$$u(x) = \sum_{j,k=1}^m \int_{-\infty}^{\infty} \varphi_j(x, \lambda) d\rho_{jk}(\lambda) \times \int_a^b \overline{\varphi_k(y, \lambda)} u(y) dy,$$

$$Pu(x) = \sum_{j,k=1}^m \int_{-\infty}^{\infty} \lambda \varphi_j(x, \lambda) d\rho_{jk}(\lambda) \times \int_a^b \overline{\varphi_k(y, \lambda)} u(y) dy,$$

where the $\rho_{jk}(\lambda)$ are functions of bounded variation and their variations represent the spectral measure. This formula shows that linear combinations of the $\varphi_j(x, \lambda)$ form generalized eigenfunctions even when λ belongs to the continuous spectrum. Later, E. C. Titchmarsh and Kodaira gave a formula to obtain the density matrix $\rho_{jk}(\lambda)$ and completed the theory (Titchmarsh [6], Kodaira [23], Dunford and Schwartz [3]). This expansion theorem makes it possible to deduce in a unified manner expansion theorems for classical special functions, such as the † Fourier series expansion theorem, the expansions by † Hermite polynomials and † Laguerre polynomials, the † Fourier integral theorem, and various expansions in terms of † Bessel functions [6, 7].

The relation between the coefficients of the differential operator and the spectral distribution of P is important in applications and is the subject of many papers [3, 5, 6].

For example, let $P(x, D) = -d^2/dx^2 + q(x)$ and $\Omega = (-\infty, \infty)$. If $q(x) \rightarrow \infty$ as $|x| \rightarrow \infty$, then $P(x, D)$ is essentially self-adjoint, the spectrum is entirely composed of the point spectrum, and a detailed estimate of the j th eigenvalue λ_j is also known [6]. If $q(x)$ converges rapidly to 0 as $|x| \rightarrow \infty$, then $P(x, D)$ is essentially self-adjoint, and there is only a continuous spectrum for $\lambda > 0$ and eigenvalues for $\lambda < 0$ with at most 0 as accumulation point. If $q(x)$ is a periodic function of x , then $P(x, D)$ is again essentially self-adjoint, and the spectrum consists of a continuous spectrum decomposed in a sequence of nonoverlapping intervals. The converse problem of determining $q(x)$ when

the spectral measure is given has been studied by I. M. Gel'fand and B. M. Levitan (*Amer. Math. Soc. Transl.*, (2) 1 (1955); original in Russian, 1951).

P. Eigenfunction Expansion for Partial Differential Operators

The theory of eigenfunction expansion for partial differential operators with continuous spectra is not complete, as it is for ordinary differential operators. What causes difficulties is that the solutions of $(P(x, D) - \lambda)u = 0$, which should be the generalized eigenfunctions, form an infinite-dimensional space, and except for special cases it is impossible to introduce convenient parameters in it. Many proofs are known for a general result, saying that any self-adjoint elliptic operator has an eigenfunction expansion into generalized eigenfunctions of the form

$$u(x) = \int_{-\infty}^{\infty} \sum_{j=1}^{k(\lambda)} \varphi_j(x, \lambda) d\rho_j(\lambda) \int_{\Omega} \overline{\varphi_j(y, \lambda)} u(y) dy,$$

[10]. There are few operators, however, for which we know how to construct $\varphi_j(x, \lambda)$ and the measure $d\rho_j(\lambda)$. The Fourier transform gives the expansion for operators with constant coefficients defined on the whole space. By means of a generalized form of the Fourier transform, Ikebe gave an eigenfunction expansion for the Schrödinger operator $-\Delta + q(x)$ in \mathbb{R}^3 under the condition that $q(x) \in L_2$ and $O(|x|^{-2-\varepsilon})$ as $|x| \rightarrow \infty$ [24]. Y. Shizuta, Mizohata, Lax and R. S. Phillips, N. A. Shenk, Ikebe, D. K. Fadeev, and others proved the same results for similar operators defined on an exterior domain with a bounded set deleted from a higher-dimensional Euclidean space. These theories are closely related to the † scattering theory of the Schrödinger equation $(-id/dt - P)u = 0$ or the wave equation $(d^2/dt^2 + P)u = 0$ associated with those operators. Lax and Phillips have developed the latter scattering theory [26] (\rightarrow 375 Scattering Theory).

Many of the problems in † quantum mechanics reduce to finding the spectral distribution of self-adjoint partial differential operators.

Q. Expansion Theorems for Non-Self-Adjoint Operators

A kind of eigenfunction expansion theorem may hold for non-self-adjoint differential operators or for non-Hilbert spaces. (See papers by the Russian school for ordinary differential operators and those by Browder,

Agmon, and others for partial differential operators.)

R. Systems of Differential Operators

We have so far dealt with single differential operators that map functions u to functions f . A linear differential operator that maps a p -tuple (u_1, \dots, u_p) of functions to a q -tuple (f_1, \dots, f_q) of functions can be written as

$$f_i(x) = \sum_{j=1}^p P_{ij}(x, D)u_j(x), \quad i = 1, \dots, q,$$

where $P(x, D) = (P_{ij}(x, D))$ is a matrix of single differential operators. Such a matrix is called a **system of differential operators**. A system $P(x, D)$ is said to be **underdetermined** if $p > q$, **determined** if $p = q$, and **overdetermined** if $p < q$. Many propositions that hold for single operators hold for determined systems under appropriate conditions.

However, there is a fundamental difference between overdetermined (underdetermined) systems and determined systems, as is seen from the theory of several complex variables, which is the theory of a typical overdetermined system $\partial/\partial\bar{z} = (\partial/\partial\bar{z}_i)$. The theory is much more difficult for overdetermined (underdetermined) systems. The general theory of overdetermined and underdetermined systems with constant coefficients has been constructed by Ehrenpreis [28], Malgrange, V. Palamodov, and Hörmander [22] for C^∞ -functions and distributions. It has been extended by H. Komatsu to the case of hyperfunctions. T. Miwa has discussed the same problem for real analytic functions.

S. Symmetric Systems of the First Order

Determined systems of the first order are important in applications. Many problems in mathematical physics are formulated in terms of them. Also, single equations of higher order can be reduced to determined systems of the first order by regarding the derivatives as unknown functions. In some cases determined systems of the first order are easier to handle than single operators of higher order. In particular, a system of differential operators

$$P(x, D) = \sum A_i(x) \partial/\partial x_i + B(x)$$

is said to be a **symmetric positive system** if the matrices $A_i(x)$ and $B(x)$ satisfy the following conditions: $A_i^* = A_i$; $B + B^* + \sum \partial A_i/\partial x_i$ is positive semidefinite. Symmetric positive systems have been studied in detail by Friedrichs [25], Phillips, C. S. Morawetz, Lax, and others.

References

- [1] K. Yosida, *Functional analysis*, Springer, 1965.
- [2] R. Courant and D. Hilbert, *Methods of mathematical physics*, Interscience, I, 1953; II, 1962.
- [3] N. Dunford and J. T. Schwartz, *Linear operators II*, Interscience, 1963.
- [4] L. Hörmander, *Linear partial differential operators*, Springer, 1963.
- [5] M. A. Naimark, *Linear differential operators I, II*, Ungar, 1967. (Original in Russian, 1954.)
- [6] E. C. Titchmarsh, *Eigenfunction expansions associated with second-order differential equations*, Clarendon Press, I, revised edition, 1962; II, 1958.
- [7] K. Yosida, *Lectures on differential and integral equations*, Interscience, 1955. (Original in Japanese, 1950.)
- [8] F. John, *Plane waves and spherical means applied to partial differential equations*, Interscience, 1955.
- [9] S. D. Eidel'man, *Parabolic systems*, North-Holland, 1969. (Original in Russian, 1964.)
- [10] I. G. Gelfand and N. Ya. Vilenkin, *Generalized functions IV, Applications of harmonic analysis*, Academic Press, 1964.
- [11] J. Leray, *Hyperbolic differential equations*, Lecture notes, Institute for Advanced Study, Princeton Univ. Press, 1953.
- [12] S. Mizohata, *Hypoellipticité des équations paraboliques*, Bull. Soc. Math. France, 85 (1957), 15–50.
- [13] J. Leray, *Problème de Cauchy I–V*, Bull. Soc. Math. France, 85 (1957), 389–429; 86 (1958), 75–96; 87 (1959), 81–180; 90 (1962), 39–156; 92 (1964), 263–361.
- [14] B. Malgrange, *Existence et approximation des solutions des équations aux dérivées partielles et des équations de convolution*, Ann. Inst. Fourier, 6 (1955–1956), 271–355.
- [15] L. Hörmander, *On the existence of real analytic solutions of partial differential equations with constant coefficients*, Inventiones Math., 21 (1973), 151–182.
- [16] I. G. Petrovskii, *Sur l'analyticité des solutions des systèmes d'équations différentielles*, Mat. Sb., 5 (47) (1939), 3–70.
- [17] L. Hörmander, *Hypoelliptic second-order differential equations*, Acta Math., 119 (1967), 147–171.
- [18] O. A. Oleinik and E. V. Radkevich, *Second-order equations with nonnegative characteristic form*, Amer. Math. Soc., 1973. (Original in Russian, 1971.)
- [19] L. Hörmander, *Seminar on singularities of solutions of linear partial differential equations*, Ann. Math. Studies 91, Princeton Univ. Press, 1979.

- [20] M. Schechter, On L^p estimates and regularity I, II, Amer. J. Math., 85 (1963), 1–13; Math. Scand., 13 (1963), 47–69.
- [21] J. J. Kohn, A priori estimates in several complex variables, Bull. Amer. Math. Soc., 70 (1964), 739–745.
- [22] L. Hörmander, An introduction to complex analysis in several variables, Van Nostrand, 1966.
- [23] K. Kodaira, On ordinary differential equations of any even order and the corresponding eigenfunction expansions, Amer. J. Math., 72 (1950), 502–544.
- [24] T. Ikebe, Eigenfunction expansions associated with the Schrödinger operators and their applications to scattering theory, Arch. Rational Mech. Anal., 5 (1960), 1–34.
- [25] K. O. Friedrichs, Symmetric positive linear differential equations, Comm. Pure Appl. Math., 11 (1958), 333–418.
- [26] P. D. Lax and R. S. Phillips, Scattering theory, Academic Press, 1967.
- [27] L. Schwartz, Théorie des distributions, Hermann, 1966.
- [28] L. Ehrenpreis, Fourier analysis in several complex variables, Wiley-Interscience, 1970.
- [29] J.-E. Björk, Rings of differential operators, North-Holland, 1979.
- [30] L. Hörmander, The analysis of linear partial differential operators, I–IV, Springer, 1983–1985.

113 (III.17) Differential Rings

Let R be a \dagger commutative ring with a unity element 1. If a map δ of R into R is such that for any pair x, y of elements of R , (i) $\delta(x+y) = \delta x + \delta y$, and (ii) $\delta(xy) = \delta x \cdot y + x \cdot \delta y$, then δ is called a **derivation** (or **differentiation**) in R . A ring R provided with a finite number of mutually commutative differentiations in R is called a **differential ring**. In this article we consider only the case where R contains a subfield that has the unity element in common with R . In particular, if R is a field, we call it a **differential field**.

In the above definition of differential ring, it is not necessary to mention the \dagger characteristic of the subfield contained in R . However, to make it more effectively applicable in the case of nonzero characteristics, we may define **differential rings using higher differentiation** in place of the differentiation defined above. If a sequence $\delta = \{\delta_v\}$ of maps $\delta_0, \delta_1, \delta_2, \dots$ of R into R satisfies the following conditions (i)–(iv) for any pair x, y of elements of R and any pair λ, μ of nonnegative integers, then δ is called a

higher differentiation in R : (i) $\delta_\lambda(x+y) = \delta_\lambda x + \delta_\lambda y$; (ii) $\delta_\lambda(xy) = \sum \delta_\alpha x \cdot \delta_\beta y$ (the addition is performed over all pairs α, β of nonnegative integers that satisfy $\alpha + \beta = \lambda$); (iii) $\delta_\lambda(\delta_\mu x) = \binom{\lambda+\mu}{\lambda} \delta_{\lambda+\mu} x$; (iv) $\delta_0 x = x$. Two higher differentiations $\delta = \{\delta_v\}$ and $\delta' = \{\delta'_v\}$ are said to be commutative if and only if δ_λ and δ'_μ commute for all pairs λ, μ of nonnegative integers. Higher differentiations were introduced by H. Hasse (1935) for the study of the field of \dagger algebraic functions of one variable in the case of nonzero characteristics.

These two definitions of differential rings coincide if the characteristic of R is zero. For simplicity, we shall restrict ourselves to that case.

Let $\delta_1, \dots, \delta_m$ be the differentiations of the differential ring R . If x is an element of R , $\delta_1^{s_1} \delta_2^{s_2} \dots \delta_m^{s_m} x(s_1, s_2, \dots, s_m$ are nonnegative integers) is called a **derivative** of x . We call x constant if and only if $\delta_1 x = \dots = \delta_m x = 0$. An \dagger ideal α of R with $\delta_i \alpha \subset \alpha$ ($i = 1, 2, \dots, m$) is called a **differential ideal** of R . If it is a \dagger prime ideal (**semiprime ideal** (i.e., an ideal containing all those elements x that satisfy $x^g \in \alpha$ for some natural number g)), then α is called a **prime differential ideal** (**semiprime differential ideal**). A subring S of R with $\delta_i S \subset S$ can be regarded as a differential ring with respect to the differentiations $\delta_1, \dots, \delta_m$. We call S a **differential subring** of R and R a **differential extension ring** of S .

Let X_1, \dots, X_n be elements of a differential extension field of a differential field K with the differentiations $\delta_1, \dots, \delta_m$, and let $\delta_1^{s_1} \delta_2^{s_2} \dots \delta_m^{s_m} X_i$ ($s_1 \geq 0, \dots, s_m \geq 0, 1 \leq i \leq n$) be \dagger algebraically independent over K . The totality of their polynomials over K , which forms a differential ring, is called the **ring of differential polynomials** of the **differential variables** X_1, \dots, X_n over K , and is denoted by $K\{X_1, \dots, X_n\}$. Its elements are called **differential polynomials**. For this ring of differential polynomials we have an analog of \dagger Hilbert's basis theorem in the ring of ordinary polynomials, **Ritt's basis theorem**: If we are given any set \mathfrak{M} of differential polynomials of X_1, \dots, X_n over K , we can choose a finite number of differential polynomials P_1, \dots, P_r from \mathfrak{M} such that each element Q of \mathfrak{M} has an integral power Q^g equal to a linear combination of P_1, \dots, P_r and their derivatives, where the coefficients of the linear combination are elements of $K\{X_1, \dots, X_n\}$. This theorem implies that in the ring of differential polynomials, every semiprime differential ideal can be expressed as the intersection of a finite number of prime differential ideals; if this expression is \dagger irredundant, it is unique (\rightarrow 67 Commutative Rings).

The equation obtained by equating a dif-

ferential polynomial to zero is called an **algebraic differential equation**. Concerning these equations, we are able to use methods similar to those used in algebraic geometry in studying the usual algebraic equations. J. Ritt made interesting studies on solutions of algebraic differential equations by such methods, principally in the case when the ground field K consists of meromorphic functions.

Since that time, basic study of differential rings and fields has been fairly well organized and has developed into theories such as the following two:

(1) **Picard-Vessiot theory**. This is a classical theory of linear homogeneous differential equations originated by E. Picard and E. Vessiot; it resembles the Galois theory concerning algebraic equations. The Galois group in this case is a linear group, and its structure characterizes the solution of the differential equation. E. Kolchin introduced the general concept of the **Picard-Vessiot extension field** of a differential field and studied in detail the group of **differential automorphisms** (i.e., the group of all those automorphisms that commute with the differentiations and fix elements of the ground field), thus making the classical theory more precise and more general.

(2) **Galois theory of differential fields**. Generalizing the concept of the Picard-Vessiot extension, Kolchin introduced the notion of the **strongly normal extension field** and established the Galois theory for such extensions. In this theory, we see that the Galois group is an algebraic group relative to a universal domain over the field of constants of the ground field. Conversely, every algebraic group is the Galois group of a strongly normal extension field. We also see that a strongly normal extension can be decomposed, in a certain sense, into a Picard-Vessiot extension and an Abelian extension (i.e., an extension whose Galois group is an Abelian variety) (\rightarrow Kolchin [2–5], Okugawa [6]).

References

- [1] I. Kaplansky, An introduction to differential algebra, Actualités Sci. Ind., Hermann, 1957.
- [2] E. R. Kolchin, Algebraic matric groups and the Picard-Vessiot theory of homogeneous ordinary linear differential equations, Ann. Math., (2) 49 (1948), 1–42.
- [3] E. R. Kolchin, Galois theory of differential fields, Amer. J. Math., 75 (1953), 753–824.
- [4] E. R. Kolchin, On the Galois theory of differential fields, Amer. J. Math., 77 (1955), 868–894.

- [5] E. R. Kolchin, Abelian extensions of differential fields, Amer. J. Math., 82 (1960), 779–790.
- [6] K. Okugawa, Basic properties of differential fields of an arbitrary characteristic and the Picard-Vessiot theory, J. Math. Kyoto Univ., 2 (1963), 295–322.
- [7] J. F. Ritt, Differential algebra, Amer. Math. Soc. Colloq. Publ., 1950.

114 (IX.18) Differential Topology

A. General Remarks

Differential topology can be defined as the study of those properties of differentiable manifolds that are invariant under diffeomorphisms. The basic objects studied in this field are the topological, combinatorial, and differentiable structures of manifolds and the relationships among them. Some of the remarkable contributions, such as H. Whitney's embedding theorem [1], the triangulation theorems of J. H. C. Whitehead and S. S. Cairnes [2, 3], and Morse theory [4], were made in the 1930s. In the late 1950s, outstanding results were obtained by R. Thom, J. W. Milnor, S. Smale, M. Kervaire, and F. Hirzebruch, among others. Differential topology thus became a new, fascinating branch of mathematics.

B. Differentiable Structures

We assume that all manifolds are paracompact. Let M be a topological manifold. A C^r -equivalence class of atlases of class C^r ($1 \leq r \leq \infty$) on M is called a **C^r -structure** on M . Any C^r -structure on M contains an atlas of class C^∞ on M , and its C^∞ -equivalence class is uniquely determined (Whitney [1]). This C^∞ -structure is called a **differentiable structure** on M compatible with the given C^r -structure. Moreover, any C^∞ -structure admits a real analytic structure compatible with it in this sense (Whitney [1]). A C^∞ -manifold is also called a **smooth manifold**, and a differentiable structure a **smooth structure**. Let D_0, D_1 be differentiable structures on M . If two C^∞ -manifolds (M, D_0) and (M, D_1) are not diffeomorphic, we say that the differentiable structures D_0, D_1 are **distinct** (\rightarrow 105 Differentiable Manifolds). Milnor defined a certain invariant of differentiable structure using the Hirzebruch index theorem (\rightarrow 56 Characteristic Classes) and proved that there are several distinct differentiable structures on the 7-dimensional sphere S^7 (Milnor [5]). Milnor's

example of such structures was: Let $f_k: S^3 \rightarrow SO(4)$ be the mapping defined by $f_k(\sigma)\tau = \sigma^h \tau \sigma^j$, where k is an odd integer, h, j are integers determined by $h+j=1$, $h-j=k$, and σ, τ are \dagger quaternions of norm 1. Let M_k^7 be the total space of the S^3 -bundle over S^4 corresponding to the mapping f_k . This is an oriented closed manifold with the naturally defined differentiable structure. Moreover, for each k , M_k^7 is homeomorphic to the 7-dimensional sphere S^7 . But if k, l are odd and $k^2 \not\equiv l^2 \pmod{7}$, M_k^7 is not diffeomorphic to M_l^7 . Differentiable manifolds (such as M_k^7) that are homeomorphic, but not diffeomorphic, to the natural sphere are called **exotic spheres**. After the discovery made by Milnor, many topological manifolds other than the 7-dimensional sphere and possessing several distinct differentiable structures have been found (N. Shimada, I. Tamura). Moreover, topological manifolds admitting no differentiable structures have been constructed (M. Kervaire [6], Smale, Tamura [7], J. Eells and N. Kuiper, C. T. C. Wall [8]). On the other hand, manifolds of dimension ≤ 3 admit unique differentiable structures.

We can introduce a \dagger Riemannian metric on a C^∞ -manifold. Let M_1^k and M_2^n be C^∞ -manifolds (k -dimensional and n -dimensional, respectively) and $f: M_1^k \rightarrow M_2^n$ be an \dagger immersion. We fix a Riemannian metric on M_2^n . For each point p of M_1^k , let $N_p(f)$ be the linear space of all tangent vectors of M_2^n at $f(p)$ that are orthogonal to the \dagger tangent space of $f(U(p))$, where $U(p)$ is a small open neighborhood of p in M_1^k . Then $\{N_p(f) | p \in M_1^k\}$ forms an $(n-k)$ -dimensional \dagger vector bundle v_f over M_1^k , called the **normal bundle** of the immersion f . The equivalence class of v_f is independent of the choice of Riemannian metric on M_2^n . When f is an \dagger embedding, we denote the submanifold $f(M_1^k)$ by L^k . Let $N_\varepsilon(L^k)$ be the set of all points whose distances (with respect to the Riemannian metric) from L^k are $\leq \varepsilon$. If L^k is compact and ε is sufficiently small, $N_\varepsilon(L^k)$ is an n -dimensional submanifold of M_2^n with boundary and is uniquely determined independently of the choice of ε up to diffeomorphism. This submanifold is called a **tubular neighborhood** of L^k in M_2^n and is denoted by $N(L^k)$. The interior of $N(L^k)$ is called an **open tubular neighborhood**. The total space $E(v_f)$ of the normal bundle v_f of the embedding f is diffeomorphic to the open tubular neighborhood of L^k in M_2^n .

C. C^r -Triangulation and the Smoothing Problem

Let (K, f) be a \dagger triangulation of the C^r -manifold M^n ($1 \leq r \leq \infty$) ($f: |K| \rightarrow M^n$) satisfying

the following conditions: (i) For any closed n -simplex σ of K , $f|_{\sigma}: \sigma \rightarrow M^n$ is a C^r -mapping; (ii) for any point p of σ , the \dagger Jacobian matrix of $f|_{\sigma}$ has rank n at p . Then we say that (K, f) is a **C^r -triangulation** of the C^r -manifold M^n , and the C^r -structure of M^n is **compatible** with the triangulation (K, f) . Concerning C^r -triangulations of C^r -manifolds, we have the following results (S. Cairns [2], J. H. C. Whitehead [3]): (i) any C^r -manifold ($1 \leq r \leq \infty$) has a C^r -triangulation, and any C^r -triangulation of the boundary can be extended to the whole manifold; (ii) for a C^r -triangulation (K, f) of a C^r -manifold M^n , the triangulated space $(K, f; M^n)$ is a \dagger combinatorial manifold; (iii) for two C^r -triangulations (K_1, f_1) and (K_2, f_2) of a C^r -manifold M^n , the triangulated spaces $(K_1, f_1; M^n)$ and $(K_2, f_2; M^n)$ are combinatorially equivalent.

Conversely, for a combinatorial manifold $(K, f; M^n)$, a differentiable structure on M that is compatible with the triangulation (K, f) is called a **smoothing** of M . The question of whether a smoothing of a combinatorial manifold exists is called the **smoothing problem**. We have some criteria for its solvability. One is expressed in terms of transverse fields, and another in terms of \dagger microbundles (\rightarrow 147 Fiber Bundles P). We also have the following method using the theory of \dagger obstruction (J. Munkres [9]; M. W. Hirsch and B. Mazur [10]). Let Γ_k be the group of oriented differentiable structures on the \dagger combinatorial k -sphere (an exact definition of this group is given in Section I). Let K^i be the i - \dagger skeleton of a combinatorial n -manifold $(K, f; M^n)$ and α a smoothing of a neighborhood U of $f(K^i)$ in M^n . Then there exists an obstruction cocycle $C_\alpha \in Z^{i+1}(K, \Gamma_i)$ satisfying the following conditions: (i) if $C_\alpha(\sigma^{i+1}) = 0$ for an $(i+1)$ -simplex σ^{i+1} of K , the smoothing α can be extended over a neighborhood of $f(K^i \cup \sigma^{i+1})$, and vice versa; (ii) if there exists another smoothing α' of a neighborhood of $f(K^i)$ that coincides with the smoothing α of a neighborhood of $f(K^{i-1})$, then $C_{\alpha'} - C_\alpha$ is a \dagger coboundary; conversely, for any $(i+1)$ -coboundary b with coefficients in Γ_i , there exists a smoothing α' of a neighborhood of $f(K^i)$ such that $b = C_{\alpha'} - C_\alpha$. Using this method we can prove that for $n \leq 7$ all combinatorial manifolds of dimension n are smoothable.

D. Embedding and Immersion Theorems

In the following, M^n and X^m are C^∞ -manifolds of dimensions n and m , respectively. Two \dagger immersions $f_0, f_1: M^n \rightarrow X^m$ are said to be **regularly homotopic** if there exists a \dagger homotopy $f_t: M^n \rightarrow X^m$, $0 \leq t \leq 1$, such that f_t is an immer-

sion for each t and the induced mapping (\dagger differential of f_t) $df_t: T(M^n) \rightarrow T(X^m)$ on the tangent spaces naturally gives rise to a continuous mapping over $T(M^n) \times I$. Two \dagger embeddings are **isotopic** if they are regularly homotopic and the homotopy f_t is an embedding for each t . Given M^n and X^m , a fundamental problem of embedding (immersion) theory is to classify the embeddings (immersions) of M^n in X^m according to their isotopy (regular homotopy) classes. Whitney proved that a continuous mapping $f: M^n \rightarrow X^m$ can be approximated by an immersion if $m \geq 2n$ and by an embedding if $m \geq 2n + 1$ [1]. The following results are also due to Whitney [1]: Any two immersions of M^n in X^m that are homotopic are regularly homotopic if $m \geq 2n + 2$, and any two embeddings of M^n in X^m that are homotopic are isotopic if $m \geq 2n + 3$. The range $m \geq 2n + 3$ is called the **stable range** of embeddings.

In [11, 12] (1944) Whitney improved his classical theorems and showed that M^n can always be immersed in \mathbf{R}^{2n-1} for $n > 1$ and M^n can always be embedded in \mathbf{R}^{2n} . The methods used in his proof have played an important role in the subsequent development. Classification of immersions of the n -sphere S^n in \mathbf{R}^m was determined by Smale [13]. Let $p: T(M^n) \rightarrow M^n$ and $p': T(X^m) \rightarrow X^m$ be the projections of the tangent bundles of M^n and X^m , respectively. A mapping $\varphi: T(M^n) \rightarrow T(X^m)$ is a **linear fiber mapping (linear fiber map)** if, for each $x \in M^n$, $\varphi(p^{-1}(x))$ is contained in a fiber of $T(X^m)$ and $\varphi|_{p^{-1}(x)}$ is a linear mapping of rank n . A **linear homotopy** $\varphi_t: T(M^n) \rightarrow T(X^m)$ ($0 \leq t \leq 1$) is a homotopy such that each φ_t is a linear fiber mapping. The following theorem of Hirsch [14] is fundamental to immersion theory: Assume $n < m$. If $\varphi: T(M^n) \rightarrow T(X^m)$ is a linear fiber mapping, the mapping $\bar{\varphi}: M^n \rightarrow X^m$ induced by φ can be approximated by an immersion $f: M^n \rightarrow X^m$ such that df and φ are linearly homotopic. Two immersions $f, g: M^n \rightarrow X^m$ are regularly homotopic if and only if df and dg are linearly homotopic. If M^n is immersible in \mathbf{R}^{m+r} , where $m > n$ with r linearly independent fields of \dagger normal vectors, then M^n is immersible in \mathbf{R}^m . In particular, if M^n is a π -manifold (\rightarrow Section I), M^n is immersible in \mathbf{R}^{n+1} .

Let $\mathcal{E}(M^n)$ be the set of isomorphism classes of real \dagger vector bundles over M^n , and consider $\theta: \mathcal{E}(M^n) \rightarrow KO(M^n)$ (\rightarrow 237 K-Theory). An element $\xi \in KO(M^n)$ is said to be **positive** if ξ is in the image of θ . If $\xi_0 \in \tilde{KO}(M^n)$, the **geometric dimension** of ξ_0 , written $g(\xi_0)$, is the least integer k such that $\xi_0 + k$ is positive. Then Hirsch's theorem [14] can be expressed as follows: M^n is immersible in \mathbf{R}^{n+k} ($k > 0$) if and only if $g(n - \tau(M^n)) \leq k$, where $\tau(M^n)$ is the tangent bundle of M^n .

A. Haefliger [15] obtained the following important result: Let M^n be compact and $\dagger(k-1)$ -connected, and let X^m be k -connected. Then any continuous mapping of M^n in X^m is homotopic to an embedding if $2k < n$, $m \geq 2n - k + 1$, and any two homotopic embeddings of M^n in X^m are isotopic if $2k < n + 1$, $m \geq 2n - k + 2$. Thus if $m > 3(n+1)/2$, any two embeddings of S^n in \mathbf{R}^m are isotopic. The range $m > 3(n+1)/2$ is called the **metastable range**. Haefliger further classified the embeddings of S^{4n-1} in \mathbf{R}^{6n} and showed the existence of embeddings of S^{4n-1} in \mathbf{R}^{6n} that are not isotopic to the natural one. More complete results for the classification of embeddings of \dagger homotopy n -spheres in S^m were obtained by J. Levine [16]. Furthermore, Levine proved the following unknotting theorem in higher-dimensional knot theory [16]: Let $f: S^{n-2} \rightarrow S^n$ be a C^∞ -embedding for $n \geq 6$; then $f(S^{n-2})$ is unknotted if and only if $S^n - f(S^{n-2})$ is homotopy equivalent to S^1 (\rightarrow 235 Knot Theory G; also 65 Combinatorial Manifolds D).

We list some results about embeddings and immersions. If M^n is noncompact, M^n can always be embedded in \mathbf{R}^{2n-1} ; if M^n is a noncompact π -manifold, M^n can always be immersed in \mathbf{R}^n ; if M^n is compact and orientable and $n > 4$, M^n can always be embedded in \mathbf{R}^{2n-1} .

E. Nonembedding and Nonimmersion Theorems

We denote the \dagger total Stiefel-Whitney class of M^n by $w(M^n)$ and the \dagger total Pontryagin class of M^n by $p(M^n)$ (\rightarrow 56 Characteristic Classes). Then $(w(M^n))^{-1} (\in H^*(M^n; \mathbf{Z}_2))$ and $(p(M^n))^{-1} (\in H^*(M^n; \mathbf{Z}))$ can be written as $\bar{w}(M^n) = \sum \bar{w}_i(M^n)$ ($\bar{w}_i \in H^i(M^n; \mathbf{Z}_2)$) and $\bar{p}(M^n) = \sum \bar{p}_i(M^n)$ ($\bar{p}_i \in H^{4i}(M^n; \mathbf{Z})$). Then the property of characteristic classes for the \dagger Whitney sum implies the following theorem: If M^n can be immersed in \mathbf{R}^{n+k} , then $\bar{w}_i(M^n) = 0$ for $i > k$ and $\bar{p}_i(M^n) = 0$ for $i > [k/2]$. Furthermore, if M^n can be embedded in \mathbf{R}^{n+k} , then $\bar{w}_k(M^n) = 0$. As an application, these results yield the nonembedding (nonimmersion) theorem for projective spaces (\rightarrow Appendix A, Table 6.VII). Sharper theorems were obtained subsequently. In particular, Atiyah proved the following: Let λ^i ($i = 0, 1, \dots$) be \dagger exterior power operations (\rightarrow 237 K-Theory), and let γ^i be the operations defined by the formal power series $\sum_{i=0}^{\infty} \gamma^i t^i = (\sum_{i=0}^{\infty} \lambda^i t^i)(1-t)^{-1}$. Then $\gamma^i(n - \tau(M^n)) = 0$ for $i > k$ ($i \geq k$) if M^n can be immersed (embedded) in \mathbf{R}^{n+k} . Furthermore, we have an interesting result for the differentiable case. For any positive integer q , there exists a differentiable

manifold M^n such that M^n is immersible in \mathbf{R}^k but not embeddable in \mathbf{R}^{k+q} .

F. Handlebodies

Let W be a C^∞ -manifold with boundary ∂W . Then the boundary ∂W has a neighborhood U that is homeomorphic to $\partial W \times \mathbf{R}_+$, where $\mathbf{R}_+ = [0, \infty)$. Let W_1, W_2 be C^∞ -manifolds with boundaries and $f: \partial W_1 \rightarrow \partial W_2$ be a diffeomorphism. Then the quotient space M of $W_1 \cup W_2$ obtained by identifying points corresponding under f has a natural differentiable structure. This construction of the C^∞ -manifold M from the C^∞ -manifolds W_1, W_2 is called **pastings together the boundaries**.

Now we consider the topological product $W_1 \times W_2$ of manifolds W_1, W_2 with boundaries $\partial W_1, \partial W_2$. $W_1 \times W_2 - \partial W_1 \times \partial W_2$ has a natural atlas of class C^∞ . By introducing a suitable atlas of class C^∞ in a neighborhood of $\partial W_1 \times \partial W_2$ in $W_1 \times W_2$, we obtain a C^∞ -manifold with boundary homeomorphic to $W_1 \times W_2$. More generally, let M be an n -manifold with boundary, let N be a finite union of submanifolds of dimension $\leq n-2$ in ∂M , and suppose that M has a corner along N . Extension of the C^∞ -structure of $M-N$ over M as in this paragraph is called **straightening the angle**.

Let D^n be the oriented unit n -disk in the n -dimensional Euclidean space \mathbf{R}^n , M_1^n and M_2^n be oriented compact C^∞ -manifolds, and $f_i: D^n \rightarrow M_i^n$ ($i=1, 2$) be C^∞ -embeddings, f_1 \dagger orientation-preserving and f_2 \dagger orientation-reversing. Then, pasting together the boundaries of $M_1^n - \text{Int } f_1(D^n)$ and $M_2^n - \text{Int } f_2(D^n)$ by the mapping $f_2 \circ f_1^{-1}$, we obtain an oriented C^∞ -manifold, called the **connected sum** of M_1^n and M_2^n and denoted by $M_1^n \# M_2^n$. The connected sum $M_1^n \# M_2^n$ has the orientation induced from those of M_1^n and M_2^n , and its differentiable structure is uniquely determined independently of the mappings f_i . Let S^n be the natural n -dimensional sphere. Then we have $M^n \# S^n \approx M^n$ (here \approx means diffeomorphic), $(M_1 \# M_2) \# M_3 \approx M_1 \# (M_2 \# M_3)$, $M_1 \# M_2 \approx M_2 \# M_1$.

Let M^n be a manifold with boundary and $f: (\partial D^s \times D^{n-s}) \rightarrow \partial M^n$ be a C^∞ -embedding. Then we call the quotient space $X(M^n; f; s)$ of $M^n \cup (D^s \times D^{n-s})$ obtained by the identification of corresponding points under f the **manifold with a handle attached by f** . Also, we call the construction of $X(M^n; f; s)$ from M^n **attaching a handle** and call $D^s \times D^{n-s}$ an s -**handle**. After straightening the angle, $X(M^n; f; s)$ is considered naturally a C^∞ -manifold with boundary. Let $f_i: \partial D_i^s \times D_i^{n-s} \rightarrow \partial M^n$ ($i=1, \dots, k$) be embeddings whose images are mutually disjoint. Then similarly, using embeddings f_1, \dots, f_k , we can define a C^∞ -manifold with handles $X(M^n;$

$f_1, \dots, f_k; s)$. In particular, $X(X(\dots(X(X(D^n; f_1; s_1); f_2; s_2) \dots); f_j; s_j))$ is called a **handlebody**.

Let M^{n-1} be an oriented $(n-1)$ -manifold and $f: \partial D^s \times D^{n-s} \rightarrow (M^{n-1} - \partial M^{n-1})$ an orientation-preserving C^∞ -embedding. Then, by straightening the angle, the quotient space of $(M^{n-1} - \text{Int } f(\partial D^s \times D^{n-s})) \cup (D^s \times \partial D^{n-s})$ obtained by the identification of points corresponding under f is a C^∞ -manifold $\chi(M^{n-1}; f)$. We say this manifold is obtained by a **spherical modification** (or **surgery**) of type $(s, n-s)$ from M^{n-1} . The manifold $\chi(M^{n-1}; f)$ has a natural orientation, and $\partial \chi(M^{n-1}; f) = \partial M^{n-1}$. The process of spherical modification has the following relation to that of attaching a handle: Let W be an n -dimensional manifold and $f: \partial D^s \times D^{n-s} \rightarrow \partial W$ an embedding. Then $\partial X(W; f; s) = \chi(\partial W; f)$. When $W = M^{n-1} \times [0, 1]$ and $f: \partial D^s \times D^{n-s} \rightarrow M \times \{1\}$, $\partial X(W; f; s) = \chi(M \times \{1\}; f) \cup M \times \{0\}$, and therefore M is cobordant (\rightarrow Section H) to $\chi(M; f)$. Conversely, let M_1 be cobordant to M_2 . Then we can obtain M_2 from M_1 by a finite sequence of spherical modifications (A. Wallace [17], Milnor). Let M^n be an n -dimensional C^∞ -manifold and $f: M^n \rightarrow \mathbf{R}^1$ be a C^∞ -function. If f satisfies the following conditions, then it is called a **nice function**: (i) All \dagger critical points of f are \dagger nondegenerate; (ii) for any critical point p of f , the \dagger index (\rightarrow 279 Morse Theory) of f at p is equal to $f(p)$. We have the following theorems.

1. Let M be a compact C^∞ -manifold. Then there exists a nice function on M (M. Morse, Smale [18], Wallace [17]).

2. Let M be a compact C^∞ -manifold and $f: M \rightarrow \mathbf{R}^1$ a C^∞ -function all of whose critical points are nondegenerate. Suppose that the number of critical points on $f^{-1}[-\varepsilon, \varepsilon]$ is k and that they are all contained in $f^{-1}(0)$. Suppose further that the indices of these critical points are all equal to s . Then $f^{-1}(-\infty, \varepsilon]$ is diffeomorphic to the manifold with handles $X(f^{-1}(-\infty, -\varepsilon]; f_1, \dots, f_k; s)$ (Morse, Thom, Smale [18]).

3. \dagger Generalized Poincaré conjecture. Let M^n be an n -dimensional \dagger homotopy sphere of class C^∞ ($n \geq 5$). Then M^n can be obtained by pasting together the boundaries of two n -disks. Consequently, M^n is homeomorphic to the n -dimensional sphere S^n (Smale [18], H. Yamaguchi [19]).

4. Let M^n be a \dagger contractible compact n -dimensional manifold ($n > 5$), with ∂M connected and simply connected. Then M^n is diffeomorphic to the n -disk D^n (Smale [18]).

5. Let M_1^n, M_2^n be oriented, compact, simply connected n -dimensional manifolds ($n > 4$). If M_1 is h -cobordant (\rightarrow Section I) to M_2 , then M_1 is diffeomorphic to M_2 (**h -cobordism theorem**; Smale [20]).

6. If $M_1^n \times \mathbf{R}^k$ is diffeomorphic to $M_2^n \times \mathbf{R}^k$, we say that M_1^n is k -**equivalent** to M_2^n . Let M_1, M_2 be compact n -dimensional manifolds of the same \dagger homotopy type. Then M_1 is k -equivalent to M_2 if and only if for a \dagger homotopy equivalence $f: M_1 \rightarrow M_2$, $\tau(M_1) \oplus \varepsilon_k$ is equivalent to $f^*\tau(M_2) \oplus \varepsilon_k$, where $\tau(M_i)$ is the \dagger tangent bundle of M_i and ε_k the \dagger trivial vector bundle of dimension k (Mazur).

7. In some special cases the classification of manifolds by diffeomorphism is completely determined. (i) The classification of simply connected 5-dimensional manifolds M with vanishing second \dagger Stiefel-Whitney class $w_2(M)$ is determined by $H_2(M)$. (ii) A 2-connected compact 6-dimensional manifold is diffeomorphic to either S^6 or the connected sum of a finite number of copies of $S^3 \times S^3$ (Smale [20]). Besides these, the classifications of $(n-1)$ -connected $2n$ -dimensional manifolds (Wall [8]) and $(n-1)$ -connected $(2n+1)$ -dimensional manifolds (Tamura [21], Wall [22]) have been obtained.

G. Thom Complexes

Let ξ be a real n -dimensional vector bundle over a paracompact space X , A_ξ be the total space of the associated bundle of ξ with fiber the closed n -disk D^n , and \dot{A}_ξ be the total space of the associated bundle of ξ with fiber the $(n-1)$ -sphere ∂D^n . Then the quotient space $X_\xi = A_\xi / \dot{A}_\xi$, obtained from A_ξ by contracting \dot{A}_ξ to a point, is called the **Thom space** of the vector bundle ξ . If X is a \dagger CW-complex, then the Thom space X_ξ of ξ has the homotopy type of a CW-complex and is called a **Thom complex**. The Thom space X_ξ has the canonical base point p_ξ corresponding to \dot{A}_ξ .

For a coefficient group G , denote by G_ξ the \dagger local system of coefficient groups with stalk G associated with the \dagger orientation sheaf of an \mathbf{R}^n -bundle ξ . Then we have the **Thom-Gysin isomorphism**:

$$H^q(X; G_\xi) \cong H^{n+q}(X_\xi, p_\xi; G),$$

$$H_q(X; G_\xi) \cong H_{n+q}(X_\xi, p_\xi; G).$$

Let G be a closed subgroup of the orthogonal group $O(n)$ and BG be the base space of the universal n -dimensional vector bundle γ_G with structure group G . Then we can take a connected CW-complex as BG . We denote the Thom complex of the vector bundle γ_G by MG and call it the **Thom complex associated with** (G, n) . The Thom complex associated with (G, n) is $(n-1)$ - \dagger connected. If G is connected, then we have $\pi_n(MG) \cong \mathbf{Z}$, $\pi_n(MO(n)) \cong \mathbf{Z}_2$, $H^n(MG; \mathbf{Z}) \cong \mathbf{Z}$, and $H^n(MO(n); \mathbf{Z}_2) \cong \mathbf{Z}_2$ (\mathbf{Z}_2 stands for the quotient group $\mathbf{Z}/2\mathbf{Z}$). The generator U of $H^n(MG; \mathbf{Z})$ is called the **fundamental (cohomology) class of the Thom complex** MG . For a general Thom space X_ξ , we can also define the fundamental class of X_ξ using the Thom-Gysin isomorphism.

mental (cohomology) class of the Thom complex MG . For a general Thom space X_ξ , we can also define the fundamental class of X_ξ using the Thom-Gysin isomorphism.

A C^∞ -submanifold W^p of a compact manifold V^n is a \dagger support of a \dagger singular cycle and represents a homology class $z \in H_p(V^n; G)$ ($G = \mathbf{Z}_2$ or \mathbf{Z}). In this case we say that the homology class z of V^n is **realizable** by a submanifold. A homology class $z \in H_{n-k}(V^n; G)$ is realizable by a submanifold if and only if there exists a mapping $f: V^n \rightarrow MO(k)$ (or $MSO(k)$) such that $f^*(U)$ is the dual cohomology class of z (Thom [23]).

The homotopy group $\pi_{n+k}(MO(n))$ (resp. $\pi_{n+k}(MSO(n))$) ($n > k$) is determined only by k for any n up to isomorphism. It is called the k -dimensional **stable homotopy group of the Thom spectrum** $MO = \{MO(n)\}$ ($MSO = \{MSO(n)\}$) and is denoted by $\pi_k(MO)$ ($\pi_k(MSO)$). For the \dagger unitary group $U(n)$ and the \dagger symplectic group $Sp(n)$, we define Thom complexes $MU(n)$ and $MSp(n)$ as the Thom complexes associated with $(U(n), 2n)$ and $(Sp(n), 4n)$ by the canonical inclusions $U(n) \subset O(2n)$ and $Sp(n) \subset O(4n)$, respectively. The k -dimensional stable homotopy groups $\pi_k(MU) = \lim \pi_{2n+k}(MU(n))$, $\pi_k(MSp) = \lim \pi_{4n+k}(MSp(n))$ of Thom spectra $MU = \{\dots, MU(n), SMU(n), MU(n+1), \dots\}$, $MSp = \{\dots, MSp(n), SMSp(n), S^2MSp(n), S^3MSp(n), MSp(n+1), \dots\}$ can be defined similarly to the case of $\pi_k(MO)$, where $SMU(n)$ denotes the \dagger reduced suspension of $MU(n)$. The stable homotopy groups of Thom spectra are calculated in connection with the cohomology groups utilizing the Thom-Gysin isomorphism (\rightarrow 202 Homotopy Theory T) (Thom [23], Milnor [24]).

H. Cobordism

Cobordism theory is a theory of classification of differentiable manifolds initiated by L. S. Pontryagin and V. A. Rokhlin, who called it **intrinsic homology**. The theory was brought to maturity by Thom [23]. Its fundamental problem is to determine whether a given compact manifold is the boundary of another manifold. Corresponding cobordism theories for combinatorial and topological manifolds are being developed (Wall, R. Williamson).

We consider only compact C^∞ -manifolds that are not necessarily connected. Let \mathfrak{D} be the set of all diffeomorphism classes of C^∞ -manifolds, and let \mathfrak{D}_0 be the set of all orientation-preserving diffeomorphism classes of oriented C^∞ -manifolds. For an oriented manifold V , we write $-V$ for the manifold with reversed orientation.

For two compact k -manifolds $V, W \in \mathfrak{D}_0$, we say that V is **cobordant** to W if there exists a compact $(k+1)$ -manifold $X \in \mathfrak{D}_0$ such that $\partial X = V \cup (-W)$. In this definition, considering \mathfrak{D} instead of \mathfrak{D}_0 , we say that V is **cobordant** to $W \bmod 2$. The equivalence class of V^k with respect to the cobordism relation (the cobordism relation mod 2) is called an **oriented (unoriented) cobordism class** and is denoted by $[V^k] ([V^k]_2)$. The set of all oriented (unoriented) cobordism classes $[V^k] ([V^k]_2)$ of k -manifolds forms an Abelian group $\Omega_k (\mathfrak{R}_k)$ by the natural addition $[V^k] + [W^k] = [V^k \cup W^k] ([V^k]_2 + [W^k]_2 = [V^k \cup W^k]_2)$. This is called the k -dimensional **oriented (unoriented) cobordism group**. Define the product $[V^k] \times [W^l] = [V^k \times W^l] ([V^k]_2 \times [W^l]_2 = [V^k \times W^l]_2)$. Then the direct sum $\Omega = \sum \Omega_k (\mathfrak{R} = \sum \mathfrak{R}_k)$ forms an anticommutative (commutative) \dagger graded algebra, which is called the **cobordism ring** or **Thom algebra**. We have the following theorems.

1. Ω_k, \mathfrak{R}_k are isomorphic to the stable homotopy groups $\pi_k(MSO), \pi_k(MO)$ of the Thom spectra MSO, MO , respectively (**Thom's fundamental theorem**) [23].

2. For a natural number k not of the form $2^i - 1$, there exists a compact manifold $P(k)$, and \mathfrak{R} is a polynomial ring over \mathbb{Z}_2 with generators $\{[P(k)]_2 | k \neq 2^i - 1\}$. $\Omega \otimes \mathbb{Q}$ (\mathbb{Q} is the field of rational numbers) is a polynomial ring over \mathbb{Q} with generators $\{[PC^{2m}] | m \geq 1\}$, where PC^{2m} is the complex $2m$ -dimensional projective space (Thom [23]). Moreover, Milnor [24] proved that the p -component of Ω_k is zero for an odd prime p . Wall proved that the 2-component of Ω_k contains no element of order 4, using a certain exact sequence which contains the natural homomorphism $\Omega_k \rightarrow \mathfrak{R}_k$ [25].

3. Let T be the ideal of Ω consisting of all torsion elements in Ω . Then Ω/T is a polynomial ring over \mathbb{Z} with generators $\{[Y_{4k}] | k \geq 1\}$, where we can take for Y_{4k} a complex $2k$ -dimensional nonsingular algebraic variety (Milnor).

\dagger Characteristic numbers are invariants of cobordism classes. Combining the results stated in 2, we have the following theorem:

4. Let V, W be manifolds. V is cobordant to $W \bmod 2$ if and only if they have the same corresponding \dagger Stiefel-Whitney numbers. Let V, W be oriented manifolds. V is cobordant to W if and only if they have the same corresponding Stiefel-Whitney numbers and \dagger Pontryagin numbers. The \dagger index of an oriented $4k$ -dimensional compact manifold is a cobordism class invariant and can be expressed as a linear combination with rational coefficients of \dagger Pontryagin numbers (\rightarrow 56 Characteristic Classes).

A manifold whose \dagger stable tangent bundle

has a complex structure is called a **stably** (or **weakly**) **almost complex manifold**. Let V, W be $2n$ -dimensional compact stably almost complex manifolds ($n \geq 1$). We say that V is **C-equivalent** to W if they have the same \dagger Chern numbers. The set of all C-equivalence classes $[V]_C$ forms the **complex cobordism group** \mathfrak{U}_n as in the (real) case of cobordism groups (the existence of the inverse of an element is not trivial). Introducing multiplication into the direct sum $\mathfrak{U}_* = \sum \mathfrak{U}_n$ by the product of manifolds, we obtain the **complex cobordism ring**. We have $\mathfrak{U}_n \cong \pi_{2n}(MU)$ and \mathfrak{U}_* is a polynomial ring over \mathbb{Z} with generators $\{[Y'_{2k}]_C | k \geq 1\}$, where we can take for Y'_{2k} a complex k -dimensional nonsingular algebraic variety (Milnor [24]).

1. h -Cobordism Groups of Homotopy Spheres

A manifold M is said to be **parallelizable** if its tangent bundle $\tau(M)$ is trivial, and **almost parallelizable** if there exists a finite number of points x_i in M such that $M - \bigcup \{x_i\}$ is parallelizable. A manifold M is called **stably parallelizable** (or **s-parallelizable**) if the \dagger Whitney sum $\tau \oplus \varepsilon_1$ of its tangent bundle $\tau(M)$ and the trivial line bundle ε_1 is trivial. A manifold M^n is called a π -**manifold** if M^n has a trivial normal bundle when it is embedded into a Euclidean space \mathbb{R}^N ($N > 2n$). A manifold M^n is a π -manifold if and only if M^n is s-parallelizable. The concepts defined in this paragraph are related by inclusions as follows: parallelizable \subsetneq s-parallelizable \subsetneq almost parallelizable. For a connected manifold with boundary, these three concepts are equivalent. \dagger Group manifolds are parallelizable. An n -dimensional manifold homeomorphic to the n -dimensional sphere is parallelizable if and only if $n = 1, 3, 7$ (Milnor [26]). Suppose that we are given an $(n-1)$ -dimensional sphere S^{n-1} (n even). We can consider the problem of determining the maximal number r such that there exists a tangent r -frame field over S^{n-1} . J. F. Adams solved this problem using \dagger K-theory (\rightarrow 237 K-Theory) as follows: Let $n = (2a+1)2^b$, $b = c+4d$, where a, b, c, d are integers and $0 \leq c \leq 3$. Put $\rho(n) = 2^c + 8d$. Then $r = \rho(n) - 1$. On the other hand, homotopy spheres are π -manifolds [27].

Let V_1, V_2 be oriented compact manifolds. Then we say that V_1 is **h -cobordant** to V_2 if there exists an oriented manifold W with boundary $\partial W = V_1 \cup (-V_2)$ such that V_i ($i = 1, 2$) is a deformation retract of W . The set of all h -cobordism classes of oriented homotopy n -spheres forms an Abelian group θ_n with connected sum as addition. This is called the (**h -cobordism**) **group of homotopy n -spheres**. We

denote by $\theta_n(\partial\pi)$ the subgroup of homotopy n -spheres that are boundaries of π -manifolds. Kervaire and Milnor gave certain exact sequences that contain the groups θ_n , the stable homotopy groups G_n of spheres, and the stable homotopy groups $\pi_n(SO)$. They clarified relations among these groups and proved that $\theta_n(\partial\pi) = 0$ for n even, $\theta_n(\partial\pi) = \text{finite group}$ for n odd $\neq 3$, $\theta_n/\theta_n(\partial\pi) \subset \text{Coker } J_n$, etc. [27], where $J_n: \pi_n(SO) \rightarrow G_n$ is the *J -homomorphism. From these results it follows that the θ_n ($n \neq 3$) are finite Abelian groups, $\theta_n = 0$ for $n < 7$, $\neq 3$, $\theta_7 \cong \mathbb{Z}_{28}$, etc. (\rightarrow Appendix A, Table 6.I).

By pasting together the boundaries of two n -disks D_1^n and D_2^n , we obtain an oriented manifold which is considered as a smoothing of the combinatorial n -sphere. The set of all orientation-preserving diffeomorphism classes of smoothings of the combinatorial n -sphere forms an Abelian group Γ_n with connected sum as addition. This is called the **group of oriented differentiable structures on the combinatorial sphere**. By the generalized Poincaré conjecture and the h -cobordism theorem, we obtain $\Gamma_n = \theta_n$ ($n \neq 3, 4$). Furthermore, $\Gamma_3 = 0$ and $\Gamma_4 = 0$ (J. Cerf [28]). The group Γ_n can also be defined as follows: Let $\text{Diff}^+ D^n$, $\text{Diff}^+ S^{n-1}$ be the **groups of orientation-preserving diffeomorphisms** of D^n , S^{n-1} , respectively, where multiplication is defined by composition. Let $r: \text{Diff}^+ D^n \rightarrow \text{Diff}^+ S^{n-1}$ be the homomorphism induced by the restriction $D^n \rightarrow S^{n-1}$. Then the image of r is a normal subgroup, and $\Gamma_n = \text{Diff}^+ S^{n-1}/r(\text{Diff}^+ D^n)$.

J. Surgery Theory

A process of modifying a manifold into another by a sequence of spherical modifications is called **surgery** on the manifold (\rightarrow Section F). The technique of surgery proved to be a powerful tool for the development of differential topology in the 1960s. Kervaire and Milnor exploited this technique in their study of homotopy spheres [27]. W. Browder [29] and S. P. Novikov [30] used surgery to construct differentiable manifolds with the same homotopy type as that of a given Poincaré complex in dimension greater than 4.

To explain the main points of surgery theory, we introduce some terminology. A pair of finite *CW complexes (X, Y) is called a **Poincaré pair of formal dimension n** if there exists a class $\mu \in H_n(X, Y; \mathbb{Z})$ called the **fundamental class** such that the \cap -product $\mu \cap: H^q(X; \mathbb{Z}) \rightarrow H_{n-q}(X, Y; \mathbb{Z})$ is an isomorphism for each q . When $Y = \emptyset$, X is called a **Poincaré complex**. Let M be an n -dimensional smooth manifold. Consider an embedding g of M into a Euclidean space \mathbb{R}^{n+k} , where k is

large enough (i.e., $k \geq n+2$). Then the isomorphism class of the *normal bundle of g is independent of the choice of embedding, and it depends only on M . Any representative of the isomorphism class is called the **normal k -vector bundle** of M and is denoted by ν_M^k . Let (X, Y) be a Poincaré pair of formal dimension n , and let ξ be a real k -vector bundle over X . A **normal mapping (normal map)** (f, b) consists of a degree 1 mapping $f: (M, \partial M) \rightarrow (X, Y)$ and a *bundle mapping $b: \nu_M^k \rightarrow \xi$ which covers f .

A normal mapping $(f, b): (M, \partial M) \rightarrow (X, Y)$ is **normally cobordant** to a normal mapping $(f', b'): (M', \partial M') \rightarrow (X, Y)$ if $\partial M = \partial M'$ and there exist a smooth $(n+1)$ -manifold W and a mapping $F: W \rightarrow X$ such that $\partial W = M \cup (-M') \cup \partial M \times I$, F is covered by a bundle mapping $B: \nu_W^k \rightarrow \xi$, and $(F, B)|_{(M, \partial M)} = (f, b)$, $(F, B)|_{(M', \partial M')} = (f', b')$.

The fundamental theorem of surgery theory is formulated as follows: Let (X, Y) be a Poincaré pair of formal dimension n . Suppose that X is simply connected and $n \geq 5$. Let $(f, b): (M, \partial M) \rightarrow (X, Y)$ be a normal mapping that restricts to a homotopy equivalence on the boundary $f|_{\partial M}: \partial M \rightarrow Y$. Then (f, b) is normally cobordant to a normal mapping $(f', b'): (M', \partial M') \rightarrow (X, Y)$ with $f': M' \rightarrow X$ a homotopy equivalence if and only if a well-defined obstruction $\sigma(f, b)$ vanishes. $\sigma(f, b)$ is called the **surgery obstruction**. When n is odd, $\sigma(f, b)$ always vanishes. When $n \equiv 0 \pmod{4}$, $\sigma(f, b)$ is an integer. If $Y = \emptyset$, it is given by $(I(M) - I(X))/8$, where $I(\quad)$ denotes the *index of the manifold or of the Poincaré complex. If $n \equiv 2 \pmod{4}$, $\sigma(f, b)$ is an integer mod 2, called the **Arf-Kervaire invariant**. For a thorough development of simply connected surgery \rightarrow [31].

In the PL or even in the topological categories, surgery theory can be developed similarly (Browder and Hirsch, R. C. Kirby and L. C. Siebenmann [32]).

In his study of Hauptvermutung for simply connected manifolds, D. Sullivan reformulated surgery in terms of the "surgery exact sequences" involving the classifying spaces G/PL or G/O . (These spaces are "homotopy theoretic fibers" of $BPL \rightarrow BG$ or $BO \rightarrow BG$ respectively.) In the special case where X is a closed simply connected PL n -manifold, the surgery exact sequence can be formulated as follows ($n \geq 5$):

$$0 \rightarrow hT(X) \xrightarrow{\eta} [X, G/PL] \xrightarrow{\theta} \begin{cases} \mathbb{Z} & (n \equiv 0 \pmod{4}), \\ \mathbb{Z}_2 & (n \equiv 2 \pmod{4}), \\ 0 & (n \text{ odd}). \end{cases}$$

The set $hT(X)$ is the totality of equivalence classes of pairs (M, f) , where M is a closed PL n -manifold and f is a homotopy equivalence $M \rightarrow X$. Two such pairs (M', f') and (M'', f'') are defined to be equivalent if there exists a PL

homeomorphism $h: M' \rightarrow M''$ such that $f'' \circ h$ is homotopic to f' . The set of homotopy classes $[X, G/PL]$ of mappings $X \rightarrow G/PL$ is appropriately identified with the set of all normal cobordism classes of normal mappings with target X or with the set of all PL reductions of the **Spivak normal fiber space** of X as a Poincaré complex [31]. The mapping η can be defined naturally, and the mapping θ corresponds to the assignment of the surgery obstruction. The image $\eta(M, f) \in [X, G/PL]$ is sometimes called the **normal invariant** of (M, f) . Sullivan reduced the classification problem of manifolds within a given homotopy type to a homotopy theoretic problem of the classifying spaces G/PL , G/O [33] (also \rightarrow [31, 34]).

In the latter half of the 1960s, surgery theory was extended by Wall to cover all compact non-simply connected manifolds which are not necessarily orientable. He introduced a certain Abelian group $L_n(\pi; w)$, now called the **Wall group**, which functorially depends on the fundamental group $\pi = \pi_1(X)$, with the orientation character $w: \pi_1(X) \rightarrow \mathbb{Z}_2$, and which is of period 4 with respect to the formal dimension n of X . In Wall's theory, the surgery obstruction $\sigma(f, b)$ takes its value in this group. The group structure of Wall groups has been calculated in many cases. Sullivan's exact sequences are extended to the non-simply connected cases [34]. Making use of the extended exact sequences, Wall and W. C. Hsiang and J. L. Shaneson classified homotopy tori. The result played an important role in the work of Kirby and Siebenmann on the π -annulus conjecture and stable homeomorphisms which led to the solution of π -Hauptvermutung and triangulation problems on topological manifolds in 1969 (\rightarrow 65 Combinatorial Manifolds). Surgery theory has many applications to other geometric problems. Among them are finding missing boundaries for open manifolds, equivariant surgery, homology surgery [35], and surgery on codimension two submanifolds [36, 37].

K. 4-Dimensional Manifolds

The results in differential topology discussed above are mainly concerned with manifolds of dimension ≥ 5 . For 4-manifolds, however, because of their peculiar nature which is not observed in other dimensions, many fundamental problems had remained unsolved until M. H. Freedman's epoch-making paper [54] appeared in 1982. His paper, together with S. K. Donaldson's theorem [55] which was published a little later, was a breakthrough in the theory of 4-manifolds.

It was Rokhlin who first discovered a strange property of 4-manifolds [38]. **Rokhlin's theorem** states: Let M be a closed oriented smooth 4-manifold. If M is almost parallelizable (or equivalently, if M is a spin 4-manifold), then the index of M is divisible by 16. Milnor and Kervaire gave an alternative proof of this theorem from the differential-topological point of view [39]. Freedman and Kirby and Y. Matsumoto gave geometric or elementary proofs.

Until 1981, Rokhlin's theorem had been a constant source of many curious phenomena in 4-dimensional topology. The list contains, for example: (1) the class $(4m+2, 4n+2) \in H_2(S^2 \times S^2) \cong \mathbb{Z} \oplus \mathbb{Z}$ cannot be represented by a smoothly embedded 2-sphere in $S^2 \times S^2$ (Kervaire and Milnor). This result was improved by W.-C. Hsiang and R. H. Szczarba, A. G. Tristram, and Rokhlin. (2) The h -cobordism theorem fails to hold in 4 or 3 dimensions (T. Matsumoto, Siebenmann [40]). (3) There exists a closed smooth 4-manifold M that is homotopy equivalent to the real projective space $\mathbb{R}P^4$, but $M \not\cong \mathbb{R}P^4$ (Cappell and Shaneson). (4) There exists an open 4-manifold W^4 properly homotopy equivalent to $S^3 \times \mathbb{R}$ but distinct (Freedman).

Closed, connected, simply connected 4-manifolds M and N are homotopy equivalent if and only if the intersection forms defined on the 2-dimensional (co)homology groups are equivalent as inner product spaces over \mathbb{Z} [41]. Wall [42] proved that if closed simply connected smooth 4-manifolds M and N are homotopy equivalent, then they are h -cobordant. Moreover, if M and N are h -cobordant, then there exists an integer $k \geq 0$ such that $M \# k(S^2 \times S^2)$ is diffeomorphic to $N \# k(S^2 \times S^2)$.

About 1973, using a certain infinite repetition process, A. Casson constructed a family of noncompact smooth 4-manifolds which are properly homotopy equivalent to the open 2-handle $D^2 \times \mathbb{R}^2$. A 4-manifold belonging to the family is called a **Casson handle**. He observed that if all the Casson handles are diffeomorphic to $D^2 \times \mathbb{R}^2$, then theories analogous to surgery and the h -cobordism theorem in higher dimensions can also be developed in dimension 4.

In his 1982 paper [54], Freedman proved that each Casson handle is homeomorphic to $D^2 \times \mathbb{R}^2$. (It was proved later that Casson handles are not, in general, diffeomorphic to $D^2 \times \mathbb{R}^2$.) This result and the proper h -cobordism theorem, also due to Freedman, proved many fundamental results on the topological structure of 4-manifolds: (1) If closed simply connected smooth 4-manifolds M , N are h -cobordant, they are homeomorphic. In

particular, a 4-dimensional homotopy 4-sphere is homeomorphic to the 4-sphere S^4 . (Proof of the 4-dimensional Poincaré conjecture.) (2) A topological 4-manifold properly homotopy equivalent to \mathbf{R}^4 is homeomorphic to \mathbf{R}^4 . (3) Given a nonsingular symmetric bilinear form ω over \mathbf{Z} , there exists a closed, connected, simply connected topological 4-manifold whose intersection form is equivalent to ω . (Therefore Rokhlin's theorem cannot be extended to topological 4-manifolds. From this, it follows that there exist simply connected topological 4-manifolds which are nonsmoothable or even nontriangulable.) (4) The homeomorphism class of a closed connected simply connected 4-manifold is determined by the intersection form and the Kirby-Siebenmann class. (This statement is an improved one by F. Quinn [56].)

Donaldson [55] revealed a sharp contrast existing between smooth and topological 4-manifolds. Donaldson's theorem states: If the intersection form of a closed, connected, simply connected smooth 4-manifold is positive definite, then the form is equivalent to the standard one $\langle 1 \rangle \oplus \langle 1 \rangle \oplus \dots \oplus \langle 1 \rangle$. This theorem, together with Casson's theory and Freedman's result (2) stated above, implies that there exists an exotic differential structure on a 4-dimensional Euclidean space \mathbf{R}^4 . Moreover, as an application of Donaldson's theorem, the problem of representing a 2-dimensional homology class of $S^2 \times S^2$ by a smoothly embedded 2-sphere was completely solved (K. Kuga [57]): the class $(p, q) \in H_2(S^2 \times S^2)$ is represented by a smoothly embedded S^2 if and only if $|p| \leq 1$ or $|q| \leq 1$.

Many problems concerning differential structures on 4-manifolds remain unsolved. It is not known whether an exotic smooth 4-sphere exists.

Among other results, the proof of the 4-dimensional annulus conjecture by Quinn [56] is remarkable.

L. Miscellaneous Results

The Browder-Livesay invariant. Let Σ^n be a homotopy n -sphere, T an **involution** on Σ^n , that is, a smooth mapping $T: \Sigma^n \rightarrow \Sigma^n$ with $T \circ T = \text{id}_{\Sigma^n}$. Assume that T is free from fixed points. The involution T is said to **desuspend** if there exists a homotopy $(n-1)$ -sphere Σ' smoothly embedded in Σ^n which is invariant under the action of T . Browder and G. R. Livesay defined an obstruction $\sigma(T)$ in the group $0(n: \text{even}), \mathbf{Z} (n \equiv 3 \pmod{4}), \mathbf{Z}_2 (n \equiv 1 \pmod{4})$ such that $\sigma(T) = 0$ if and only if T desuspends (provided that $n \geq 6$) [43]. The invariant $\sigma(T)$ is now called the **Browder-**

Livesay invariant. The technique used there was equivariant surgery under the action of T .

De Rham homotopy theory. A new approach to topology of manifolds was invented by Sullivan [44], who considered the exterior algebra of differential forms (with polynomial coefficients) on a simplicial complex. Through the construction of the **minimal model** for the algebra of differential forms, he recovers the rational homotopy type of the complex under some reasonable condition on the fundamental group. Following upon the classical de Rham theorem which recovers the (real) cohomology algebra from the algebra of differential forms, this approach is called the **de Rham homotopy theory**. This method has proved to be useful for the explicit calculation of the algebraic topological invariants of $^*\text{Kähler}$ manifolds, $^*\text{loop}$ spaces, cross-section spaces, $^*\text{algebraic}$ varieties, etc.

Kirby calculus. Kirby [45] initiated a link-theoretic approach to 3- and 4-manifolds. Let L be a link in S^3 . Suppose that each component of L is given a **framing**, which means a trivialization of a tubular neighborhood. Such a link is called a **framed link**. A framing of a component is determined by the $^*\text{linking}$ number between the component and its replacement along the framing. By attaching 2-handles to the 4-disk D^4 along a framed link in $S^3 = \partial D^4$, we obtain a handlebody; here we denote it by W_L . It is known that each closed connected oriented 3-manifold is obtained as the boundary of such a handlebody (W. B. R. Lickorish, Wallace). Kirby proved that for framed links L and L' , the boundaries ∂W_L and $\partial W_{L'}$ belong to the same orientation-preserving diffeomorphism class if and only if L is transformed into L' by a sequence of the following two kinds of elementary operations: (i) adding or deleting a trivial knot (separated from L) with framing ± 1 , (ii) band summation of components corresponding to "handle sliding." Kirby's approach is called **Kirby calculus** on framed links. For applications \rightarrow [45, 53].

References

- [1] H. Whitney, Differentiable manifolds, *Ann. Math.*, (2) 37 (1936), 645–680.
- [2] S. S. Cairns, Triangulation of the manifold of class one, *Bull. Amer. Math. Soc.*, 41 (1935), 549–552.
- [3] J. H. C. Whitehead, On C^1 -complexes, *Ann. Math.*, (2) 41 (1940), 809–824.
- [4] M. Morse, The calculus of variations in the large, *Amer. Math. Soc. Colloq. Publ.*, 1934.
- [5] J. W. Milnor, On manifolds homeomor-

- phic to the 7-sphere, *Ann. Math.*, (2) 64 (1956), 399–405.
- [6] M. A. Kervaire, A manifold which does not admit any differentiable structure, *Comment. Math. Helv.*, 34 (1960), 257–270.
- [7] I. Tamura, 8-manifolds admitting no differentiable structure, *J. Math. Soc. Japan*, 13 (1961), 377–382.
- [8] C. T. C. Wall, Classification of $(n-1)$ -connected $2n$ -manifolds, *Ann. Math.*, (2) 75 (1962), 163–189.
- [9] J. Munkres, Obstructions to the smoothing of piecewise differentiable homeomorphisms, *Ann. Math.*, (2) 72 (1960), 521–554.
- [10] M. W. Hirsch and B. Mazur, *Smoothings of piecewise linear manifolds*, *Ann. Math. Studies*, Princeton Univ. Press, 1974.
- [11] H. Whitney, The self-intersections of a smooth n -manifold in $2n$ -space, *Ann. Math.*, (2) 45 (1944), 220–246.
- [12] H. Whitney, The singularities of a smooth n -manifold in $(2n-1)$ -space, *Ann. Math.*, (2) 45 (1944), 247–293.
- [13] S. Smale, The classification of immersions of spheres in Euclidean spaces, *Ann. Math.*, (2) 69 (1959), 327–344.
- [14] M. W. Hirsch, Immersions of manifolds, *Trans. Amer. Math. Soc.*, 93 (1959), 242–276.
- [15] A. Haefliger, Plongements différentiables de variétés dans variétés, *Comment. Math. Helv.*, 36 (1961), 47–82.
- [16] J. Levine, A classification of differentiable knots, *Ann. Math.*, (2) 82 (1965), 15–50.
- [17] A. H. Wallace, Modifications and cobounding manifolds, *Canad. J. Math.*, 12 (1960), 503–528.
- [18] S. Smale, Generalized Poincaré's conjecture in dimensions greater than four, *Ann. Math.*, (2) 74 (1961), 391–406.
- [19] H. Yamasage, On the Poincaré conjecture for M^5 , *J. Math. Osaka City Univ.*, 12 (1961), 1–17.
- [20] S. Smale, On the structure of manifolds, *Amer. J. Math.*, 84 (1962), 387–399.
- [21] I. Tamura, Classification des variétés différentiables, $(n-1)$ -connexes, sans torsion, de dimension $2n+1$, *Sém. H. Cartan* 16–19, 1962–1963, *Inst. H. Poincaré*, Univ. Paris, 1964.
- [22] C. T. C. Wall, Classification problems in differential topology, I, II, *Topology*, 2 (1963), 253–261, 263–272.
- [23] R. Thom, Quelques propriétés globales des variétés différentiables, *Comment. Math. Helv.*, 28 (1954), 17–86.
- [24] J. W. Milnor, On the cobordism ring Ω^* and a complex analogue I, *Amer. J. Math.*, 82 (1960), 505–521.
- [25] C. T. C. Wall, Determination of the cobordism ring, *Ann. Math.*, (2) 72 (1960), 292–311.
- [26] J. W. Milnor, Some consequences of a theorem of Bott, *Ann. Math.*, (2) 68 (1958), 444–449.
- [27] M. A. Kervaire and J. W. Milnor, Groups of homotopy spheres I, *Ann. Math.*, (2) 77 (1963), 504–537.
- [28] J. Cerf, Sur les difféomorphismes de la sphère de dimension trois ($\Gamma_4=0$), *Lecture notes in math.* 53, Springer, 1968.
- [29] W. Browder, Homotopy type of differentiable manifolds, *Colloquium on Algebraic Topology*, Aarhus, 1962, 42–46.
- [30] S. P. Novikov, Diffeomorphisms of simply connected manifolds (in Russian), *Dokl. Akad. Nauk. SSSR*, 143 (1962), 1046–1049.
- [31] W. Browder, Surgery on simply connected manifolds, *Erg. Math.*, 65 (1972).
- [32] R. C. Kirby and L. C. Siebenmann, *Foundational essays on topological manifolds, smoothings, and triangulations*, *Ann. Math. Studies* 88, Princeton Univ. Press, 1977.
- [33] D. Sullivan, Triangulating and smoothing homotopy equivalences and homeomorphisms, *Geometric Topology Seminar Notes*, Princeton Univ. Press, 1967.
- [34] C. T. C. Wall, *Surgery on compact manifolds*, Academic Press, 1970.
- [35] S. E. Cappell and J. L. Shaneson, The codimension two placement problem and homology equivalent manifolds, *Ann. Math.*, (2) 99 (1974), 277–348.
- [36] Y. Matsumoto, Knot cobordism groups and surgery in codimension two, *J. Fac. Sci. Univ. Tokyo*, 20 (1973), 253–317.
- [37] M. H. Freedman, Surgery on codimension 2 submanifolds, *Mem. Amer. Math. Soc.*, vol. 12, no. 191, 1977.
- [38] V. A. Rohlin (Rohlin), New results in the theory of 4-dimensional manifolds (in Russian), *Dokl. Akad. Nauk SSSR*, 84 (1952), 221–224.
- [39] J. W. Milnor and M. A. Kervaire, Bernoulli numbers, homotopy groups, and a theorem of Rohlin, *Proc. Int. Congress Math.*, 1958, Cambridge Univ. Press, 1960, 454–458.
- [40] L. C. Siebenmann, Disruption of low-dimensional handle body theory by Rohlin's theorem, *Topology of Manifolds*, Markham, 1970, 57–76.
- [41] J. W. Milnor, On simply connected 4-manifolds, *Symp. Int. de Top. Alg.* (Mexico 1956), Mexico, 1958, 122–128.
- [42] C. T. C. Wall, On simply connected 4-manifolds, *J. London Math. Soc.*, 39 (1964), 141–149.
- [43] S. López de Medrano, Involutions on manifolds, *Erg. Math.*, 59, 1971.
- [44] D. Sullivan, Infinitesimal computation in topology, *Publ. Math. Inst. HES*, 47 (1977), 269–331.
- [45] R. C. Kirby, A calculus for framed links in S^3 , *Inventiones Math.*, 45 (1978), 35–56.

- [46] J. Munkres, Elementary differential topology, Ann. Math. Studies, Princeton Univ. Press, 1961.
- [47] S. Smale, A survey of some recent developments in differential topology, Bull. Amer. Math. Soc., 69 (1963), 131–145.
- [48] C. T. C. Wall, Topology of smooth manifolds, J. London Math. Soc., 40 (1965), 1–20.
- [49] J. W. Milnor, Lectures on the h -cobordism theorem, Princeton Univ. Press, 1965.
- [50] R. Stong, Notes on cobordism theory, Princeton Univ. Press, 1968.
- [51] J. W. Milnor, Topology from the differentiable viewpoint, Univ. Press of Virginia, 1969.
- [52] M. W. Hirsch, Differential topology, Springer, 1976.
- [53] R. Mandelbaum, Four-dimensional topology: An introduction, Bull. Amer. Math. Soc., (New Series) 2 (1980), 1–159.
- [54] M. H. Freedman, The topology of four-dimensional manifolds, J. Differential Geometry, 17 (1982), 357–453.
- [55] S. K. Donaldson, An application of gauge theory to four-dimensional topology, J. Differential Geometry, 18 (1983), 279–315.
- [56] F. Quinn, Ends of maps III: dimension 4 and 5, J. Differential Geometry, 17 (1982), 503–521.
- [57] K. Kuga, Representing homology classes of $S^2 \times S^2$, Topology, 23 (1984), 133–137.

115 (XVII.8) Diffusion Processes

A. General Remarks

Let $(\Omega, \mathfrak{B}, P)$ be a * probability space. A * Markov process $\{X_t\}_{0 \leq t < \infty}$ on a * topological space S with * continuous time parameter t is called a **diffusion process** if the * sample function $X_t(\omega)$ is continuous in t with probability 1 until a random time $\zeta(\omega)$, called the * terminal time. After the terminal time, $X_t(\omega)$ stays at the terminal point ∂ . Such a process is said to be **1-dimensional** or **multidimensional** according as S is an interval or a manifold (possibly with boundary) with dimension ≥ 2 . Brownian motion is a typical diffusion process (\rightarrow 45 Brownian Motion).

To give conditions for a Markov process to be a diffusion process, let us assume that S is a * complete metric space with metric ρ . Let $\{X_t\}$ be a Markov process on S with time parameter t ranging over a finite interval $[t_1, t_2]$ and satisfying

$$\sup P(s, x, t, S - U_\varepsilon(x)) = O(h) \quad (h \downarrow 0) \quad (1)$$

for every $\varepsilon > 0$, where $U_\varepsilon(x)$ is the ε -neighbor-

hood of x and sup is taken over all $x \in S$, and $s, t \in [t_1, t_2]$ such that $0 < t - s < h$. Then $\{X_t\}$ is a diffusion process if and only if

$$\int_{t_1}^{t_2} P(\rho(X_t, X_{t+h}) > \varepsilon) dt = o(h) \quad (h \downarrow 0)$$

for every $\varepsilon > 0$ [5]. This includes the following result due to E. B. Dynkin (1952) and J. R. Kinney (1953) as a special case: If in (1) we can replace $O(h)$ with $o(h)$, then $\{X_t\}$ will be a diffusion process. Specifically, when $\{X_t\}$ is a 1-dimensional * strong Markov process, the latter condition is also necessary for the process to be a diffusion process.

Diffusion processes are intimately related to a certain class of * partial differential equations. Let S be the real line. Assume that the * transition probability $P(s, x, t, E)$ ($s < t$) of the process $\{X_t\}_{0 \leq t < \infty}$ satisfies

$$1 - P(s, x, s + h, (x - \varepsilon, x + \varepsilon)) = o(h) \quad (h \downarrow 0) \quad (2a)$$

for every $\varepsilon > 0$ and that the following limits exist:

$$\lim_{h \rightarrow 0+} \frac{1}{h} \int_{x-\varepsilon}^{x+\varepsilon} (y-x)^2 P(s, x, s+h, dy) = 2a(s, x) > 0, \quad (2b)$$

$$\lim_{h \rightarrow 0+} \frac{1}{h} \int_{x-\varepsilon}^{x+\varepsilon} (y-x) P(s, x, s+h, dy) = b(s, x), \quad (2c)$$

$$\lim_{h \rightarrow 0+} \frac{1}{h} (P(s, x, s+h, S) - 1) = c(s, x) \leq 0. \quad (2d)$$

Assume further that the transition probability is * absolutely continuous with respect to Lebesgue measure: $P(s, x, t, dy) = p(s, x, t, y) dy$. Then under some suitable additional conditions, $p(s, x, t, y)$ satisfies

$$\frac{\partial p}{\partial s} = -a(s, x) \frac{\partial^2 p}{\partial x^2} - b(s, x) \frac{\partial p}{\partial x} - c(s, x)p, \quad (3)$$

and

$$\frac{\partial p}{\partial t} = \frac{\partial^2}{\partial y^2} (a(t, y)p) - \frac{\partial}{\partial y} (b(t, y)p) + c(t, y)p, \quad (4)$$

where δ is the * Dirac delta function. The coefficient c vanishes if $\{X_t\}$ is * conservative. Equations (3) and (4) are called **Kolmogorov's backward equation** and **forward equation**, respectively. They are also called the **Fokker-Planck partial differential equations**.

A. N. Kolmogorov [1] derived equations (3) and (4) in 1931, and W. Feller proved that (3) (or (4)) has a unique solution under certain regularity conditions on the coefficients a , b , and c , and that the solution $p(s, x, t, y)$ is non-

negative, has an integral with respect to y that does not exceed 1, and satisfies the †Chapman-Kolmogorov equation

$$\int p(s, x, t, y) p(t, y, u, z) dy = p(s, x, u, z)$$

for every $0 < s < t < u$. Hence $p(s, x, t, y)$ determines a †Markov process analytically. However, rigorous proof establishing the sufficiency of condition (2) for the Markov process $\{X_t\}$ to be a diffusion process did not appear until the 1950s.

In the †temporally homogeneous case, $p(s, x, s+h, y)$ does not depend on s and can be written as $p(h, x, y)$. Then $a(s, x)$, $b(s, x)$, and $c(s, x)$ are also independent of s , and $p(t, x, y)$ satisfies

$$\frac{\partial p}{\partial t} = a(x) \frac{\partial^2 p}{\partial x^2} + b(x) \frac{\partial p}{\partial x} + c(x)p,$$

$$p(0+, x, y) = \delta(x - y). \quad (5)$$

Feller made an intensive study of this case and completely solved the problem of existence and uniqueness of the solution of (5) assuming that $p(t, x, y)$ is nonnegative and that its integral with respect to y does not exceed 1 [7]. In particular, when t varies in $[0, +\infty)$ and S is an interval $[r_1, r_2]$, Feller used the Hille-Yosida theory of †semigroups of operators to determine the conditions that should be satisfied by r_1 and r_2 in order that the differential equation (5) (with the initial condition and his additional assumptions) yield one and only one solution. Feller also introduced the notion of generalized differential operators, which expresses the differential operator in the right-hand side of (5) in the most general form [8]. The probabilistic meaning of his results was clarified by Dynkin, H. P. McKean, K. Itô, D. B. Ray, and others, and all 1-dimensional diffusion processes with the †strong Markov property have now been completely determined.

Since not much research on temporally nonhomogeneous diffusion processes has been done so far, we restrict our explanation to temporally homogeneous ones. Let $\mathfrak{M} = (X_t, W, P_x | x \in S)$ be a Markov process, where S is a †state space, W is the †path space consisting of all paths $w: [0, +\infty) \rightarrow S \cup \{\partial\}$ which are continuous in t for $0 \leq t < \zeta(w)$ ($w(t) = \partial$ for $t \geq \zeta(w)$ while $w(t) \in S$ for $0 \leq t < \zeta(w)$), and P_x is a probability measure on W under the condition that the process starts from x at $t = 0$ (\rightarrow 261 Markov Processes). We can actually identify W with the †basic space Ω and set $X_t(w) = w(t)$ for $w \in \Omega$. Assume that \mathfrak{M} has the †strong Markov property. It follows from †Dynkin's formula for †infinitesimal generators (\rightarrow 261 Markov Processes) that the infinitesimal generator \mathfrak{G} of a diffusion process has the local property stating that if u and v belong to the domain of \mathfrak{G} and coincide in a neighborhood of x_0 , then $\mathfrak{G}u(x_0) = \mathfrak{G}v(x_0)$.

B. 1-Dimensional Diffusion Processes

Let S be a straight line. A point $x \in S$ is called a **right singular point** if $X_t(w) \geq x$ for all $t \in [0, \zeta(w))$ with P_x -probability 1. A **left singular point** is defined analogously, with \geq replaced by \leq . A right and left singular point is called a **trap**, while a right singular point which is not left singular is called a **right shunt** (a **left shunt** is defined analogously). A point is called a **regular point** if it is neither right nor left singular.

The set of all regular points is open. Let (r_1, r_2) be a connected component of this open set. One of the most important results concerning this situation is the proof of the existence of a strictly increasing function $s(x)$ defined on (r_1, r_2) and two measures m and k on (r_1, r_2) such that the infinitesimal generator \mathfrak{G} of \mathfrak{M} is represented as

$$\mathfrak{G}u(x) = \frac{u^+(dx) - u(x)k(dx)}{m(dx)}, \quad (6)$$

where $u^+(dx)$ is the measure $du^+(x)$ induced by the †right derivative $u^+(x)$ of $u(x)$ with respect to $s(x)$ (i.e., $u^+(x) = \lim_{\Delta x \rightarrow +0} \{u(x + \Delta x) - u(x)\} / \{s(x + \Delta x) - s(x)\}$). Equation (6) gives a generalization of second-order †differential operator $au'' + bu' + cu$ ($a > 0, c \leq 0$) [12]. Here m is positive for nonempty open sets, both m and k are finite for compact sets in (r_1, r_2) , and s, m , and k are unique in the following sense: If there are two sets of values of s_i, m_i , and k_i ($i = 1, 2$), then $s_2(x) = cs_1(x) + \text{constant}$, $m_2(dx) = c^{-1}m_1(dx)$, and $k_2(dx) = c^{-1}k_1(dx)$ for some positive constant c . We call s, m , and k , respectively, the **canonical scale**, **canonical measure** (or **speed measure**), and **killing measure** for \mathfrak{M} . They determine the behavior of $X_t(w)$ belonging to \mathfrak{M} inside the interval (r_1, r_2) . Conversely, given any such set of s, m , and k , we can find a 1-dimensional diffusion process \mathfrak{M} such that s, m , and k are, respectively, the canonical scale, canonical measure, and killing measure of \mathfrak{M} . If $X_t(w)$ is nonvanishing in (r_1, r_2) with probability 1, the killing measure k is identically zero, and the canonical scale s satisfies the equation

$$P_x(\sigma_{x_2} < \sigma_{x_1}) = \frac{s(x) - s(x_1)}{s(x_2) - s(x_1)}$$

for $x_1 < x < x_2$, where σ_y is the †hitting time of the point y .

The motion $X_t(w)$ belonging to the process \mathfrak{M} and contained in (r_1, r_2) can be constructed

from the Brownian motion by means of a topological transformation of the state space (interval) based on s , a †time change based on m , and a †killing based on k . More precisely, we first transform the interval (r_1, r_2) by $x \rightarrow s(x)$ into the interval $(s(r_1 + 0), s(r_2 - 0))$ so that the diffusion process on this new interval has a canonical scale coincident with x . The speed and killing measures are transformed accordingly. We can, therefore, assume that the canonical scale is x . Let us consider the case $(r_1, r_2) = (-\infty, +\infty)$ for simplicity. Let $t(t, x)$ be the †local time of Brownian motion at x (\rightarrow 45 Brownian Motion). Next, we apply the †time change to the Brownian motion by means of the †additive functional

$$\varphi(t) = \int_{-\infty}^{+\infty} t(t, x) m(dx),$$

and finally †kill the latter process by means of the †multiplicative functional

$$\alpha(t) = \exp\left(-\int_{-\infty}^{+\infty} t(\varphi^{-1}(t), x) k(dx)\right).$$

Thus we obtain the process \mathfrak{M} in $(-\infty, \infty)$. In particular, if $m(dx) = a(x)^{-1} dx$ and $k(dx) = |c(x)| dx$, we have

$$\varphi(t) = \int_0^t a(X_\tau)^{-1} d\tau,$$

$$\alpha(t) = \exp\left(\int_0^t c(X_{\varphi^{-1}(\tau)}) d\tau\right),$$

and $\mathfrak{G}u = au'' + cu$.

At a shunt the infinitesimal generator \mathfrak{G} has a form that is a generalization of the first-order differential operator $bu' + cu$ ($c \leq 0$), with $b > 0$ or $b < 0$ according as it is a right shunt or a left shunt. At a trap we have $\mathfrak{G}u(x) = -u(x)/E_x(\zeta)$.

When S is an interval with endpoints r_1 and r_2 and all interior points are regular, the left endpoint r_1 is classified into the following 4 types, according to the behavior of \mathfrak{M} near r_1 : Take an arbitrary fixed point $r \in (r_1, r_2)$, and set $n(dx) = m(dx) + k(dx)$ and

$$\alpha = \int_{(r_1, r)} (s(r) - s(x)) n(dx),$$

$$\beta = \int_{(r_1, r)} n((x, r)) s(dx).$$

Then r_1 is a **regular boundary** if $\alpha < \infty$, $\beta < \infty$; an **entrance boundary** if $\alpha < \infty$, $\beta = \infty$; an **exit boundary** if $\alpha = \infty$, $\beta < \infty$; and a **natural boundary** if $\alpha = \infty$, $\beta = \infty$. This classification is independent of the choice of r . A similar classification of r_2 can be established. $X_t(w)$ approaches r_1 in finite time with positive or null probability according as β is finite or infinite. If $\alpha = \infty$, it never happens that $X_t(w)$

starts from r_1 and reaches the interior of the interval S even if $r_1 \in S$, whereas if $\alpha < \infty$ we can construct (adjoining r_1 to S if necessary) a diffusion process that enters the interior from r_1 and whose motion in the interior coincides with that of $X_t(w)$.

If r_1 is a regular boundary for \mathfrak{M} and $r_1 \in S$, then there are various possibilities for the behavior of $X_t(w)$ at r_1 . They are expressed by the boundary conditions satisfied by the functions u belonging to the domain of the infinitesimal generator \mathfrak{G} . The condition is in general of the form

$$\gamma u(r_1) + \delta \mathfrak{G}u(r_1) + \mu u^+(r_1) = 0,$$

where γ , δ , and μ are constants, $\gamma, \delta \leq 0$, $\mu \geq 0$, and $|\delta| + \mu > 0$. If $\gamma = \delta = 0$, then r_1 is said to be a **reflecting barrier**. If r_1 is regular for \mathfrak{M} and does not belong to S , then $X_t(w)$ vanishes exactly as $X_t(w)$ reaches r_1 , and r_1 is called an **absorbing barrier**. This case corresponds to the boundary condition $u(r_1) = 0$. Whatever the boundary condition may be, \mathfrak{M} is constructed from the Brownian motion with reflecting barrier by topological transformation of the state space, time change, and killing. Here if $\gamma \neq 0$, then killing may occur at r_1 ; if $\delta \neq 0$, the set of visiting times of r_1 has positive Lebesgue measure; and if $\mu \neq 0$, the trace of the motion may go beyond the point r_1 and reach the interior points of S [2, 7–9].

If we weaken the assumption of continuity of paths and admit jumps from r_1 , the general boundary condition becomes

$$\gamma u(r_1) + \delta \mathfrak{G}u(r_1) + \mu u^+(r_1) + \int_{(r_1, r_2]} (u(x) - u(r_1)) v(dx) = 0,$$

where v is a measure with respect to which $\min(1, s(x) - s(r_1 + 0))$ is integrable.

When $S = (r_1, r_2)$, the transition probability is absolutely continuous with respect to the canonical measure, the density $p(t, x, y)$ has an †eigenfunction expansion

$$p(t, x, y) = \int_{-\infty}^0 e^{\lambda t} e(d\lambda; x, y),$$

and $p(t, x, y)$ is positive, jointly continuous in 3 variables, and symmetric in x and y . A similar result is also known when S is half-open or closed [9].

If \mathfrak{M} is †recurrent, i.e., $P_x(\sigma_y < +\infty) = 1$ for every x and y in S , then there exists a unique (up to a multiplicative constant) †invariant measure for S that is finite for all closed intervals in the interior of S . If \mathfrak{M} is conservative and all the interior points of S are regular, then the canonical measure is an invariant measure, provided that the endpoints are either entrance, natural, or regular reflecting.

C. Multidimensional Diffusion Processes

Let the state space S be a domain or the closure of a domain in the n -dimensional Euclidean space \mathbf{R}^n . Consider a temporally homogeneous diffusion process $\{X_t\}_{0 \leq t < \infty}$ on S . Under suitable regularity conditions the infinitesimal generator \mathfrak{G} coincides, for sufficiently smooth functions u in its domain, with the following \dagger elliptic partial differential operator A :

$$A = \sum_{i,j=1}^n a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^n b^i(x) \frac{\partial}{\partial x^i} + c(x), \quad c \leq 0; \quad (7)$$

where S has a boundary, u satisfies a boundary condition of the form

$$\sum_{i,j=1}^n \alpha^{ij}(x) \frac{\partial^2 u(x)}{\partial x^i \partial x^j} + \sum_{i=1}^{n-1} \beta^i(x) \frac{\partial u(x)}{\partial x^i} + \gamma(x)u(x) + \delta(x)Au(x) + \mu(x) \frac{\partial u(x)}{\partial n} = 0, \quad (8)$$

where, for simplicity, we assume that S is the closed half-space defined by $x^n \geq 0$, (α^{ij}) is a symmetric nonnegative definite matrix, $\gamma \leq 0$, $\delta \leq 0$, $\mu \geq 0$, and $\partial/\partial n$ is the inward-directed \dagger conormal derivative associated with a^{ij} . This boundary condition was discovered by A. D. Venttsel' [13]. Conversely, given an operator A such as (7) and a boundary condition (8) (if S has a boundary), the existence and uniqueness of the corresponding diffusion process are known for several special cases. If $S = \mathbf{R}^n$ and A has continuous coefficients, there exists at least one diffusion process corresponding to A [14–16].

A probabilistic approach to constructing diffusion processes corresponding to the operator A with $c=0$ is given by Itô's method of \dagger stochastic differential equations (\rightarrow 406 Stochastic Differential Equations E). A somewhat different approach was introduced by D. W. Stroock and S. R. S. Varadhan [20] under the name of **martingale problems** (\rightarrow 261 Markov Processes C). Let $W^n = C([0, \infty) \rightarrow \mathbf{R}^n)$ be the space of all continuous functions $w: [0, \infty) \rightarrow \mathbf{R}^n$ endowed with compact uniform topology and $\mathfrak{B}(W^n)$ be the topological σ -field. Given $x \in \mathbf{R}^n$, a **solution to the martingale problem** for the operator A with $c=0$ starting from x is a probability measure P_x on $(W^n, \mathfrak{B}(W^n))$ satisfying $P_x(w(0)=x) = 1$ such that $f(w(t)) - \int_0^t Af(w(s))ds$ is a P_x -martingale for all $f \in C_0^\infty(\mathbf{R}^n)$, where $C_0^\infty(\mathbf{R}^n)$ denotes the set of all C^∞ -functions on \mathbf{R}^n having compact support. If $a = (a^{ij})$ is uniformly positive definite, bounded, and continuous and if $b = (b^i)$ is bounded and Borel measurable, the martingale problem for the operator A with $c=0$ is **well posed**, i.e., for each $x \in \mathbf{R}^n$, there is exactly

one solution starting from x . (For further information related to the theory of martingale problems \rightarrow [15, 20].)

Suppose next that S is the closure of a bounded domain with a sufficiently smooth boundary and A is given with sufficiently smooth coefficients. If the boundary condition is $\gamma u + \mu \partial u / \partial n = 0$, $\mu \neq 0$, then there exists a unique diffusion process on S corresponding to this situation. Moreover, if $\gamma = 0$, then the process is said to have a **reflecting barrier**. S. Watanabe [22] gave a probabilistic condition which characterizes the reflecting diffusion processes in the normal direction among all reflecting diffusion processes in oblique directions. Suppose that a general boundary condition (8) is given. We write the left-hand side of (8) as Lu . Write $u = Hf$ for the solution of $Au = 0$ with boundary value f . Under some natural additional conditions, T. Ueno proved that if $LH \dagger$ generates a Markov process on the boundary, then there exists a diffusion process for A with boundary condition (8) [21]. If $\{X_t\}$ has a reflecting barrier, the Markov process on the boundary is, conversely, obtained from $\{X_t\}$ through time change by a nonnegative continuous \dagger additive functional which increases only when the value of X_t is on the boundary. Stochastic differential equations are also used in constructing diffusion processes with boundary condition (8) [19] (\rightarrow 406 Stochastic Differential Equations). An elegant method for constructing diffusion processes with boundary condition (8) has been introduced by Watanabe [22]. It consists of piecing together excursions from the boundary to the boundary. In this construction, the stochastic integrals for \dagger Poisson point processes play an important role. There are various other results on multidimensional diffusion processes with general boundary conditions which cannot be covered by the method mentioned above (\rightarrow [23] and M. Motoo, *Proc. Intern. Symp. Stochastic Differential Equations, Kyoto, 1976*).

When we have a diffusion process on S with infinitesimal generator of the form A , we can obtain a probabilistic expression for the solutions of various partial differential equations involving A . Let σ be the hitting time of the boundary of S . Then $Hf(x)$ can be expressed as $E_x(f(X_{\sigma-0}))$. The solution of $Au = -f$ with boundary value 0 is given by $u(x) = E_x(\int_0^\sigma f(X_t)dt)$, while the solution $u(t, x)$ of $\partial u / \partial t = Au$ with boundary value 0 and initial value $u(0, x) = f$ is $E_x(f(X_t); t < \sigma)$. The first case gives the solution of a \dagger Dirichlet problem; and in this case, the condition for a boundary point to be \dagger regular relative to the Dirichlet problem can also be expressed probabilistically (\rightarrow 45 Brownian Motion, 261 Markov

Processes). Furthermore, if $f(X_t)$ is replaced by $f(X_t)\exp(\int_0^t k(X_s)ds)$ in the expressions for $u(x)$ and $u(t, x)$, A is replaced by $A + k$ (M. Kac [24]). When $k \leq 0$, this replacement gives rise to a killing of the process (\rightarrow 261 Markov Processes E).

By using the theory of \dagger Dirichlet forms we can investigate a general class of \dagger symmetric multidimensional diffusion processes which are not in the framework of the classical diffusion processes, i.e., diffusion processes whose infinitesimal generators are not necessarily differential operators (M. Fukushima, *Dirichlet Forms and Markov Processes*, 1980)).

D. Diffusion Processes on Manifolds

Let M be a \dagger connected \dagger oriented $\dagger\sigma$ -compact C^∞ \dagger manifold of dimension n . Let \hat{M} be M or $M \cup \{\partial\}$ (= the one-point \dagger compactification of M) according as M is \dagger compact or non-compact. Suppose that we are given a system of C^∞ \dagger vector fields A_0, A_1, \dots, A_r on M . We consider the following stochastic differential equation on M :

$$dX_t = \sum_{k=1}^r A_k(X_t) \circ dw^k(t) + A_0(X_t)dt \quad (9)$$

(\rightarrow 406 Stochastic Differential Equations), where $w(t) = (w^1(t), w^2(t), \dots, w^r(t))$ denotes an r -dimensional Brownian motion and the first term of the right-hand side is understood in the sense of the \dagger Stratonovich stochastic differential. Let $X(t, x, w)$ be the solution of (9) with $X_0 = x \in M$ defined on the r -dimensional Wiener space (W_0^r, P^w) (\rightarrow 406 Stochastic Differential Equations) and P_x be the probability law on $\hat{W}(M)$ of $(X(t, x, w))_{t \geq 0}$, where $\hat{W}(M)$ is the space of all continuous mappings $w: [0, \infty) \rightarrow \hat{M}$ with ∂ as a trap. Then $\{P_x | x \in M\}$ defines a diffusion process \mathfrak{M} on M which is generated by the second-order differential operator $\sum_{k=1}^r A_k^2/2 + A_0$, i.e., a diffusion process \mathfrak{M} with the infinitesimal generator \mathfrak{G} such that

$$\mathfrak{G}f(x) = \frac{1}{2} \sum_{k=1}^r A_k^2 f(x) + A_0 f(x) \quad \text{for } f \in C_0^\infty(M),$$

where $C_0^\infty(M)$ denotes the space of all C^∞ -functions on M with compact support. By appealing to the analytical theory of partial differential equations we can discuss regularity properties of the transition probability of \mathfrak{M} ([15]; L. Hörmander, *Acta Math.*, 119 (1967)). Recently, P. Malliavin [25] also suggested a probabilistic method for proving elliptic regularity results (see also [19]).

We now assume that all linear sums of A_0, A_1, \dots, A_r are \dagger complete. Then the terminal

time of \mathfrak{M} is infinite a.s. Let $W(M)$ be the space of all continuous mappings $w: [0, \infty) \rightarrow M$ endowed with compact uniform topology, and set $W_x(M) = \{w | w \in W(M), w(0) = x\}$. We describe the \dagger topological support $\mathcal{S}(P_x)$ of the probability P_x on $W_x(M)$, i.e., the smallest closed subset of $W_x(M)$ that carries the measure P_x . Let \mathcal{S} be the set of all piecewise constant mappings $u: [0, \infty) \rightarrow \mathbf{R}^r$. For a given $u = (u^1(t), u^2(t), \dots, u^r(t))$ of \mathcal{S} , we consider a system of ordinary differential equations

$$\frac{d}{dt} \varphi(t) = A_0(\varphi(t)) + \sum_{k=1}^r A_k(\varphi(t))u^k(t), \quad (10)$$

i.e., for every C^∞ -function f on M with compact support,

$$\frac{d}{dt} f(\varphi(t)) = A_0 f(\varphi(t)) + \sum_{k=1}^r (A_k f)(\varphi(t))u^k(t).$$

Then for every $u \in \mathcal{S}$ and $x \in M$, we obtain a curve $\varphi = \varphi(x, u) = (\varphi_t(x, u))$ on M by solving (10) with $\varphi(0) = x$. Set $\mathcal{S}^x = \{\varphi(x, u) | u \in \mathcal{S}\}$. Then we have

$$\mathcal{S}(P_x) = \overline{\mathcal{S}^x} \quad \text{for every } x \in M.$$

Let \mathcal{L} be the \dagger Lie algebra generated by A_1, A_2, \dots, A_r , and set $\mathcal{L}(x) = \{V_x | V \in \mathcal{L}\}$. If $\dim \mathcal{L}(x) = n$ for every $x \in M$, then $\mathcal{S}(P_x) = W_x(M)$ for every $x \in M$. For further information \rightarrow Stroock and Varadhan (*Proc. 6th Berkeley Symp. Math. Statist. Prob.* III, 1972) and H. Kunita (*Proc. Int. Symp. Stochastic Differential Equations, Kyoto*, 1976).

Let A be a smooth \dagger nondegenerate second-order \dagger elliptic differential operator on M which is expressed in local coordinates as

$$A = \frac{1}{2} \sum_{i,j=1}^n a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^n \bar{b}^i(x) \frac{\partial}{\partial x^i},$$

where $(a^{ij}(x))$ is symmetric and \dagger strictly positive definite. Then there exist a \dagger Riemannian metric g and a C^∞ \dagger vector field b on M such that

$$A = \frac{1}{2} \Delta_M + b, \quad (11)$$

where Δ_M is the \dagger Laplace-Beltrami operator on the \dagger Riemannian manifold (M, g) . We now construct the diffusion process generated by the operator A , introducing a stochastic differential equation on the \dagger bundle $O(M)$ of the \dagger orthonormal frames. There exists an \dagger affine connection ∇ \dagger compatible with the Riemannian metric g such that for every C^∞ -function f on M ,

$$\frac{1}{2} \sum_{k=1}^n L_k^2(f \circ \pi)(r) = \left(\frac{1}{2} \Delta_M + b \right) (f)(\pi(r)),$$

$$r \in O(M),$$

where (L_1, L_2, \dots, L_n) is the system of \dagger canonical horizontal vector fields (\dagger basic vector fields) on $O(M)$ corresponding to the affine connection ∇ and $\pi: O(M) \rightarrow M$ is the natural projection [19]. We now consider the following stochastic differential equation on $O(M)$:

$$dr_t = \sum_{k=1}^n L_k(r_t) \circ dw^k(t). \quad (12)$$

Let $r(t, r, w)$ be the solution of (12) with $r_0 = r \in O(M)$ defined on the n -dimensional Wiener space (W_0^n, P^W) . Now a stochastic curve $X(t, r, w)$ on M is defined by $X(t, r, w) = \pi(r(t, r, w))$. Set $X(t, r, w) = \partial$ for $t \geq \zeta$, where ζ is the \dagger explosion time of $r(t, r, w)$. Then the probability law of $(X(t, r, w))_{t \geq 0}$ on $\hat{W}(M)$ depends only on $x = \pi(r)$, and it defines a diffusion process \mathfrak{M} on M which is generated by the operator A of (11). (For details \rightarrow [19, 26].) When $A = \Delta_M/2$, i.e., $b = 0$ in (11), the diffusion process \mathfrak{M} is called the \dagger Brownian motion on the Riemannian manifold M (\rightarrow 45 Brownian Motion).

Next consider the case when (L_1, L_2, \dots, L_n) in (12) is a system of canonical horizontal vector fields corresponding to the \dagger Riemannian connection, and let b be a C^∞ -vector field on M . Let \tilde{b} be the scalization of b , i.e., $\tilde{b} = (\tilde{b}^1, \tilde{b}^2, \dots, \tilde{b}^n)$, $\tilde{b}^i(r) = \sum_{j=1}^n b^j(x) f_j^i$, and $(f_j^i) = (e_j^i)^{-1}$ for $r = (x, e_1, e_2, \dots, e_n) \in O(M)$, where $b = \sum_{i=1}^n b^i \partial / \partial x^i$ and $e_i = \sum_{j=1}^n e_j^i \partial / \partial x^j$ in local coordinates. Suppose that $\tilde{b}^i(r)$ is bounded on $O(M)$ for $i = 1, 2, \dots, n$. Then

$$M(t) = \exp \left[\sum_{i=1}^n \int_0^t \tilde{b}^i(r_s) dw^i(s) - \frac{1}{2} \sum_{i=1}^n \int_0^t (\tilde{b}^i(r_s))^2 ds \right]$$

is an \dagger exponential martingale. The diffusion process generated by the operator $\Delta_M/2 + b$ is obtained from the Brownian motion on M constructed above by the \dagger transformation of drift, i.e., the transformation by the \dagger multiplicative functional $M(t)$ (\rightarrow 261 Markov Processes). (G. Maruyama (*Nat. Sci. Rep. Ochanomizu Univ.*, 15 (1954)) studied this for 1-dimensional processes, I. V. Girsanov [27] and M. Motoo (*Ann. Inst. Statist. Math.*, 12 (1960–1961)) for multidimensional cases.)

Consider a diffusion process $\mathfrak{M} = \{P_x | x \in M\}$ on M generated by the operator A of (11). Let ω_b be the \dagger differential 1-form defined from the vector field b by

$$\omega_b(V)_x = g_x(b_x, V_x)$$

for every $x \in M$ and every C^∞ -vector field V on M . Then \mathfrak{M} is symmetrizable if and only if ω_b is \dagger exact, i.e., if there exists a function F on M such that $\omega_b = dF$. The \dagger invariant measures

are then of the form $\text{const.} \times \exp[2F(x)]m(dx)$, where $m(dx)$ is the \dagger Riemannian volume (E. Nelson, *Duke Math. J.*, 25 (1958); Kolmogorov, *Math. Ann.*, 113 (1937)). The diffusion process \mathfrak{M} is said to be **locally symmetrizable** if for every \dagger simply connected domain $D \subset M$ there exists a Borel measure $\nu^D(dx)$ on D such that

$$\int_D T_t^D f(x) g(x) \nu^D(dx) = \int_D f(x) T_t^D g(x) \nu^D(dx)$$

for all bounded continuous functions f and g , where

$$T_t^D f(x) = \int_{\hat{W}(M)} f(w(t)) I_{\{t | \tau_D(w) > t\}} P_x(dw),$$

$$\tau_D(w) = \inf \{t | w(t) \notin D\}.$$

Then \mathfrak{M} is locally symmetrizable if and only if ω_b is \dagger closed, i.e., $d\omega_b = 0$. R. Z. Khas'minskiĭ [28] proved a pair of useful tests for explosions of diffusion processes on $M = \mathbf{R}^n$, similar to Feller's test for $n = 1$ mentioned in Section B (\rightarrow [18]). In case of a general manifold we can investigate the possibility of explosions for \mathfrak{M} by appealing to the comparison theorems for \dagger curvatures in the theory of differential geometry [19, 29]; S. T. Yau, *J. Math. Pures Appl.*, 57 (1978)). The diffusion process \mathfrak{M} on M is said to be \dagger recurrent if $P_x[X_t \in U \text{ for some } t > 0] = 1$ for any open subset U of M ; otherwise it is called \dagger transient. It is well known that an n -dimensional Brownian motion on \mathbf{R}^n is recurrent if $n \leq 2$ and transient if $n \geq 3$ [9]. There also are some results for the criterion which determines whether \mathfrak{M} is recurrent or transient ([19, 28]; A. Friedman, *Stochastic Differential Equations and Applications* I, II, 1975, and K. Ichihara, *Publ. Res. Inst. Math. Sci.*, 14 (1978)). Explicit formulas for the transition probability of the Brownian motion on a hyperbolic space are given in [29]. For information related to the asymptotic behavior of diffusion processes as $t \downarrow 0$ we refer the reader to Varadhan (*Comm. Pure Appl. Math.*, 20 (1967)) and S. A. Molchanov (*Russian Math. Surveys*, 30 (1975)).

Diffusion approximations for suitably normalized random sequences have been studied extensively beginning with the work of A. J. Khinchin [30]. The theory of \dagger limiting distributions of sums of \dagger independent (or weakly dependent) random variables is among the best-known examples (\rightarrow 250 Limit Theorems in Probability Theory). For information related to the theory of diffusion approximations, see Yu. V. Prokhorov (*Theory of Prob. Appl.*, 1 (1956)), A. V. Skorokhod [16], and Stroock, Varadhan, and G. Papanicolaou (*Proc. 1976 Duke Turbulence Conference*). Recently, many interesting examples of multidimensional diffusion processes have also been

introduced to describe probabilistic models in physics, biology, etc. (e.g., R. L. Stratonovich, *Topics in the Theory of Random Noise* I, II, 1963; J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory*, 1970; and K. Sato, *Proc. Int. Symp. Stochastic Differential Equations*, Kyoto, 1976).

In general, if a diffusion process $\{X_t\}$ is given on a noncompact space S and X_t has no limit points in S as $t \uparrow \zeta$, then some natural compactification should be induced by $\{X_t\}$. The notion of a * Martin boundary for Markov processes is introduced in this connection.

References

- [1] A. N. Kolmogorov, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Math. Ann.*, 104 (1931), 451–458.
- [2] E. B. Dynkin, *Markov processes* I, II, Springer, 1965. (Original in Russian, 1963.)
- [3] J. L. Doob, *Stochastic processes*, Wiley, 1953.
- [4] K. Itô, *Lectures on stochastic processes*, Tata Inst., 1960.
- [5] L. V. Seregin, Continuity conditions for stochastic processes, *Theory of Prob. Appl.*, 6 (1961), 1–26. (Original in Russian, 1961.)
- [6] W. Feller, Zur Theorie der Stochastischen Prozesse (Existenz- und Eindeutigkeits-satz), *Math. Ann.*, 113 (1936), 113–160.
- [7] W. Feller, The parabolic differential equations and the associated semi-groups of transformations, *Ann. Math.*, (2) 55 (1952), 468–519.
- [8] W. Feller, On second order differential operators, *Ann. Math.*, (2), 61 (1955), 90–105.
- [9] K. Itô and H. P. McKean, Jr., *Diffusion processes and their sample paths*, Springer, 1965.
- [10] P. Lévy, *Processus stochastiques et mouvement brownien*, Gauthier-Villars, 1948.
- [11] E. B. Dynkin, One-dimensional continuous strong Markov processes, *Theory of Prob. Appl.*, 4 (1959), 1–52. (Original in Russian, 1959.)
- [12] I. S. Kats (Kac) and M. G. Kreĭn, On the spectral functions of the string, *Amer. Math. Soc. Transl.*, (2) 103 (1974). (Original in Russian, supplement of the Russian translation of the book by F. V. Atkinson, *Discrete and continuous boundary problems*.)
- [13] A. D. Venttsel' (Ventcel'), On boundary conditions for multidimensional diffusion processes, *Theory of Prob. Appl.*, 4 (1959), 164–177. (Original in Russian, 1959.)
- [14] N. V. Krylov, On the selection of a Markov process from a system of processes and the construction of quasidiffusion processes, *Math. USSR-Izv.*, 7 (1973), 691–709. (Original in Russian, 1973.)
- [15] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, Springer, 1979.
- [16] A. V. Skorokhod, *Studies in the theory of random processes*, Addison-Wesley, 1965. (Original in Russian, 1961.)
- [17] K. Itô, On stochastic differential equations, *Mem. Amer. Math. Soc.*, 1951.
- [18] H. P. McKean, Jr., *Stochastic integrals*, Academic Press, 1969.
- [19] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland and Kodansha, 1981.
- [20] D. W. Stroock and S. R. S. Varadhan, Diffusion processes with continuous coefficients I, II, *Comm. Pure. Appl. Math.*, 22 (1969), 345–400, 479–530.
- [21] K. Sato and T. Ueno, Multidimensional diffusions and the Markov process on the boundary, *J. Math. Kyoto Univ.*, 4 (1965), 529–605.
- [22] S. Watanabe, Poisson point process of Brownian excursions and its applications to diffusion processes, *Amer. Math. Soc. Proc. Symp. Pure Math.*, 31 (1977), 153–164.
- [23] M. Motoo, Application of additive functionals to the boundary problem of Markov processes, *Proc. 5th Berkeley Symp. Math. Stat. Prob.* II, pt. 2, Univ. of California Press, 1967, 75–110.
- [24] M. Kac, On some connections between probability theory and differential and integral equations, *Proc. 2nd Berkeley Symp. Math. Stat. Prob.*, Univ. of California Press, 1951, 189–215.
- [25] P. Malliavin, C^k -hypoelliptic with degeneracy, *Stochastic Analysis*, A. Friedman and M. Pinsky (eds.), Academic Press, 1978, 199–214, 327–340.
- [26] P. Malliavin, Formule de la moyenne, calcul de perturbations et théorèmes d'annulation pour les formes harmoniques, *J. Functional Anal.*, 17 (1974), 274–291.
- [27] I. V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, *Theory of Prob. Appl.*, 5 (1960), 285–301. (Original in Russian, 1960.)
- [28] R. Z. Khas'minskii, Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations, *Theory of Prob. Appl.*, (1960), 179–196. (Original in Russian, 1960.)
- [29] A. Debiard, B. Gaveau and E. Mazet, Théorèmes de comparaison en géométrie riemannienne, *Publ. Res. Inst. Math. Sci.*, 12 (1976), 391–425.
- [30] A. Ya. Khinchin, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Springer, 1933.

116 (XX.2) Dimensional Analysis

The system of units for physical quantities is derived from a certain set of **fundamental units**. If the fundamental units are denoted by θ, φ, ψ , etc., any other unit α (called a **derived unit**) can always be expressed in the form $\alpha = c\theta^l\varphi^m\psi^n \dots$ (c, l, m, n, \dots are constants), by definition or by physical laws. The exponents l, m, n, \dots are called the **dimensions** of α , and the content of the previous statement is expressed as $[\alpha] = [\theta^l\varphi^m\psi^n \dots]$, which is called the **dimensional formula**. The usual practice is to take as fundamental units length, time, mass, temperature, and energy, which are denoted by L, T, M, θ , and H , respectively. **Dimensional analysis** investigates the relation between physical quantities by use of the π theorem and the law of similitude given below.

A. The π Theorem

If a relationship $f(\alpha, \beta, \dots) = 0$ holds among n physical quantities α, β, \dots independently of the choice of fundamental units, the equation $f(\alpha, \beta, \dots) = 0$ can always be transformed into $F(\pi_1, \pi_2, \dots) = 0$, where the π_i are $n - m$ dimensionless quantities (m is the number of fundamental units) of the form $\pi_i = \alpha^{a_i}\beta^{b_i} \dots$. If we choose the π_i so that $\pi_1 = \alpha\beta^{-r_1}\gamma^{-z_1} \dots$ and π_2, π_3 , etc. do not contain α , then $f = 0$ implies $\alpha = \beta^{r_1}\gamma^{z_1} \dots \Phi(\pi_2, \pi_3, \dots)$, which clearly shows the manner in which the quantity α is related to other quantities β, γ, \dots .

B. The Law of Similitude

In general, if two physical systems of the same kind have the same values of the π_i , then the physical states of the systems are similar. If we are given a family of mutually similar systems, it is sufficient to observe a particular one among them (a "model") in order to estimate physical values attached to any one of the given systems.

Consider, for example, the case of the drag D acting on geometrically similar bodies placed in the flow of a viscous incompressible fluid. If v is the velocity, l the representative length of the body, ρ the density of the fluid, and μ the viscosity (which has the dimensional formula $ML^{-1}T^{-1}$), then the π theorem gives $D/\rho v^2 l^2 = f(\rho v l / \mu)$. Hence the drag coefficient as given by the left-hand side can be obtained by the experiments performed on a geometrically similar model. The dimensionless quantity $R = \rho v l / \mu$ ($v = \mu / \rho$) is called the **Reynolds number**. If the wave resistance due to gravity

as well as the effect of compressibility are taken into account, gravitational acceleration g and sound velocity a must be included, so that we have $D/\rho v^2 l^2 = f(vl/v, v^2/lg, v/a, C_1, C_2, \dots)$, where C_1, C_2, \dots are other dimensionless quantities depending on the physical properties of the fluid. $Fr = v^2/lg$ is called the **Froude number**, and $M = v/a$ the **Mach number**.

Next, consider the case of heat transfer between a solid surface and a flowing fluid. Let the area of the solid surface be denoted by S , the heat transferred per unit time by Q (HT^{-1}), the thermal conductivity of the fluid by k ($HL^{-1}T^{-1}\theta^{-1}$), the specific heat by C ($HM^{-1}\theta^{-1}$), the two representative temperatures by T_0 and T_1 , and the representative length by l , where expressions in parentheses represent dimensional formulas. Then we have, as dimensionless quantities, the **Nusselt number** $Nu = Q/(kS(T_1 - T_0)/l)$, the **Prandtl number** $Pr = \nu/\kappa$ ($\kappa = k/\rho c$), the **Grashoff number** $Gr = d^3 g(T_1 - T_0)/\nu^2 T_0$, and R , so that from the π theorem we have the relation $Nu = f(R, Pr, Gr, C_1, C_2, \dots)$. Furthermore, $Pe = \nu l / \kappa = Pr R$ is called the **Péclet number**.

References

- [1] A. W. Porter, The method of dimensions, Methuen, third edition, 1946.
 - [2] D. C. Ipsen, Units, dimensions, and dimensionless numbers, McGraw-Hill, 1960.
 - [3] H. L. Langhaar, Dimensional analysis and theory of models, Wiley, 1951.
 - [4] H. E. Huntley, Dimensional analysis, Macdonald, 1952.
 - [5] C. M. Focken, Dimensional methods and their applications, Arnold, 1953.
 - [6] R. Kurth, Dimensional analysis and group theory in astrophysics, Pergamon, 1972.
 - [7] L. I. Sedov, Similarity and dimensional methods in mechanics, Academic Press, 1959. (Original in Russian, fourth edition, 1957.)
- Also → 414 Systems of Units.

117 (II.21) Dimension Theory

A. Introduction

Toward the end of the 19th century, G. Cantor discovered that there exists a one-to-one correspondence between the set of points on a line segment and the set of points on a square; and also, G. Peano discovered the existence of a †continuous mapping from the segment onto the square. Soon, the progress of the theory of

point-set topology led to the consideration of sets which are more complicated than familiar sets, such as polygons and polyhedra. Thus it became necessary to give a precise definition to dimension, a concept which had previously been used only vaguely. In 1913, L. E. J. Brouwer [9] gave a definition of dimension based on an idea of H. Poincaré. In 1922, the foundations of dimension theory for separable metric spaces were established by K. Menger [11] and P. Uryson [10]. Subsequently, P. S. Aleksandrov and W. Hurewicz contributed much to the development of the theory. The foundations of dimension theory for general metric spaces were established independently by M. Katětov and K. Morita. More general theory for \dagger normal spaces has also been investigated; the same results as in metric spaces, however, do not always hold.

B. Definition of Dimension

Let X be a normal space. If any finite open \dagger covering of X has an open covering of \dagger order $\leq n + 1$ as its refinement (\rightarrow 425 Topological Spaces R) (i.e., if for any open sets G_i ($i = 1, \dots, s$) such that $X = G_1 \cup \dots \cup G_s$, there exist open sets H_i ($i = 1, \dots, s$) such that $H_i \subset G_i$, $X = H_1 \cup \dots \cup H_s$ and any $n + 2$ of the H_i have no point in common), then we write $\dim X \leq n$. If $\dim X \leq n$ but $\dim X \leq n - 1$ does not hold, then we define X to be n -dimensional and write $\dim X = n$. We call $\dim X$ the **covering dimension**, (or **Lebesgue dimension**) of X . The idea behind this definition is due to H. Lebesgue.

There are other definitions of *dimension* that are given inductively. Let us define $\text{Ind } X = -1$ if X is empty. If for any pair consisting of a closed set F and an open set G with $F \subset G$ in X there exists an open set V such that $F \subset V \subset G$ and $\text{Ind}(\bar{V} - V) \leq n - 1$, then we define $\text{Ind } X \leq n$. Next, we define $\text{ind } \emptyset = -1$. For any point p of X and any neighborhood G of p , suppose that there exists an open neighborhood V of p such that $V \subset G$ and $\text{ind}(\bar{V} - V) = n - 1$. Then we define $\text{ind } X \leq n$. As before, we set $\text{Ind } X = n$ ($\text{ind } X = n$) if $\text{Ind } X \leq n$ ($\text{ind } X \leq n$) but $\text{Ind } X \leq n - 1$ ($\text{ind } X \leq n - 1$) does not hold. (The definition of $\text{ind } X$ is due to Menger.) We call $\text{Ind } X$ ($\text{ind } X$) the **large inductive dimension** of X (the **small inductive dimension** of X).

If $\dim X \leq n$ does not hold for any n , then X is called **infinite-dimensional**, written $\dim X = \infty$; we define $\text{Ind } X = \infty$ and $\text{ind } X = \infty$ similarly. These dimensions are invariant under \dagger homeomorphisms.

The set of irrational points in a Euclidean

space, the \dagger Cantor discontinuum, and \dagger Baire zero-dimensional spaces are all 0-dimensional. The set of rational points in a separable \dagger Hilbert space is 1-dimensional.

C. Dimension of Metric Spaces

The following theorems hold for the dimension of metric spaces (M. Katětov, *Czechoslovak Math. J.*, 2 (1952); K. Morita, *Math. Ann.*, 128 (1954)). Let X and Y be metric spaces. The equality $\dim X = \text{Ind } X$ holds. If $Y \subset X$, then $\dim Y \leq \dim X$. If X is a union of a countable number of closed sets F_i ($i = 1, 2, \dots$), then $\dim X = \max(\dim F_i)$ (**sum theorem for dimension**). The inequality $\dim(X \cup Y) \leq \dim X + \dim Y + 1$ holds. If $\dim X = n$, then X is a union of $n + 1$ 0-dimensional subsets (**decomposition theorem for dimension**). We have $\dim(X \times Y) \leq \dim X + \dim Y$, where $X \neq \emptyset$ (**product theorem for dimension**).

Each of the following is a necessary and sufficient condition for $\dim X \leq n$: (i) There exists a subspace A of a Baire zero-dimensional space B (τ) and a continuous closed mapping f of A onto X such that $f^{-1}(x)$ consists of at most $n + 1$ points for each point x of X (K. Morita, *Sci. Rep. Tokyo Kyoiku Daigaku*, 5 (1955)); (ii) there exists a metric of X which gives the same topology on X such that for any positive number ε , any point x of X , and any $n + 2$ points x_i ($i = 1, \dots, n + 2$) at a distance less than ε from the $(\varepsilon/2)$ -neighborhood of x , there are at least two points x_i and x_j ($i \neq j$) with distance $< \varepsilon$ (J. Nagata, *Fund. Math.*, 45 (1958)).

Hurewicz's problem asked whether the equality $\dim X = n + m$ ($m > 0$) implies the existence of an m -dimensional space A and a mapping f of A onto X having property (i). It was solved affirmatively for separable metric spaces by J. H. Roberts and for general metric spaces by K. Nagami (*Japan. J. Math.*, 30 (1960)).

If X is the union of a countable number of closed \dagger strongly paracompact subspaces, in particular if X is separable, then $\text{Ind } X = \text{ind } X$ [1, 2]. However, it was shown by P. Roy (*Bull. Amer. Math. Soc.*, 68 (1962)) that this equality does not hold in general.

D. Euclidean Spaces and Dimension

The n -dimensional \dagger Euclidean space \mathbf{R}^n is exactly n -dimensional in the sense mentioned above; thus this concept of dimension agrees with our intuition. The proof of $\dim \mathbf{R}^n \geq n$ comes from **Lebesgue's theorem**: If each member of a finite closed covering of an n -cube has

sufficiently small diameter, then the order of the covering is not less than $n + 1$. (The proof of $\dim \mathbf{R}^n \leq n$ is easy.) Let X be a subset of \mathbf{R}^n and f a homeomorphism from X onto a subset $f(X)$ of \mathbf{R}^n . If x is an interior point of X , then $f(x)$ is an interior point of $f(X)$. Also, if an open set A of \mathbf{R}^n is homeomorphic to a subset B of \mathbf{R}^n , then B is open in \mathbf{R}^n (**Brouwer's theorem on the invariance of domain** [8]).

This theorem holds for any manifold but not for general separable metric spaces. By the theorem of invariance of domain it can be shown that \mathbf{R}^m and \mathbf{R}^n , $m \neq n$, are not homeomorphic (**theorem on invariance of dimension of Euclidean spaces**). Any n -dimensional separable metric space is embedded in a Euclidean space \mathbf{R}^{2n+1} , or, more precisely, in the subset of \mathbf{R}^{2n+1} consisting of all points x of which at most n coordinates are rational (**Menger-Nöbeling embedding theorem**, G. Nöbeling, *Math. Ann.*, 104 (1930)). Thus, from the topological point of view, any finite-dimensional separable metric space can be identified with a subset of a Euclidean space. Moreover, it is known that any n -dimensional separable metric space is homeomorphic to a subset of some n -dimensional compact metric space.

If F is a bounded closed subset of \mathbf{R}^m , then $\dim F \leq n$ if and only if for any positive number ε , there exists a continuous mapping f from F into an n -dimensional polyhedron in \mathbf{R}^m such that the distance between x and $f(x)$ is less than ε for each point x of F .

E. Dimension of Normal Spaces

Let X be a normal space. Then $\text{Ind } X \geq \dim X$ and $\text{Ind } X \geq \text{ind } X$, but the equalities do not necessarily hold here. The following theorems were obtained by E. Čech, Aleksandrov, C. H. Dowker, E. Hemmingsen, and Morita [1]. If $\dim X \leq n$, then any locally finite open covering of X has an open covering of order $\leq n + 1$ as its refinement; if A is an ${}^{\dagger}F_{\sigma}$ subset of X or A is strongly paracompact, then $\dim A \leq \dim X$; if X has a ${}^{\dagger}\sigma$ -locally finite closed covering $\{F_{\alpha}\}$, then $\dim X = \max(\dim F_{\alpha})$.

In order that $\dim X \leq n$, it is necessary and sufficient that any continuous mapping from a closed subset of X into an n -sphere S^n can be extended continuously to X . If X and Y are † paracompact and X is † locally compact, or if $X \times Y$ is strongly paracompact, then $\dim(X \times Y) \leq \dim X + \dim Y$, where $X \neq \emptyset$; if X is a † CW complex, then the equality holds [14]. M. L. Wage (*Proc. Nat. Acad. Sci. US*, 75 (1978)) proved under the continuum hypothesis (CH) that $\dim(X \times Y) \leq \dim X + \dim Y$

does not hold in general even if $X \times Y$ is locally compact and normal, and $\dim X = \dim Y = 0$; T. Przymusiński (*Proc. Amer. Math. Soc.*, 76 (1979)) noted that CH can be avoided by a modification of Wage's construction. Katětov (*Časopis Pěst. Mat. Fys.*, 75 (1950)) proved that $\dim X$ is determined by the ring $C^*(X)$ of bounded real-valued continuous functions on X .

F. Homological Dimension

Aleksandrov contributed much to the development of dimension theory in introducing the concept of homological dimension (*Math. Ann.*, 106 (1932)). The **homological dimension** of a compact Hausdorff space X with respect to an Abelian group G is the largest integer n such that the n -dimensional † Čech homology group $\check{H}_n(X, A; G)$ is nonzero for some closed subset A of X . The **cohomological dimension** $D(X; G)$ is defined similarly by using the † Čech cohomology group $\check{H}^n(X, A; G)$. If $\dim X < \infty$, then $\dim X = D(X; \mathbf{Z})$ (\mathbf{Z} is the additive group of integers). The cohomological dimension of X with respect to an arbitrary Abelian group is determined by the cohomological dimension with respect to some specified groups, and the cohomological dimension of the product space $X \times Y$ is expressed in terms of those of X and Y (M. F. Bokshtein, [5]). A compact Hausdorff space X has the property that $\dim(X \times Y) = \dim X + \dim Y$ for any compact Hausdorff space Y if and only if $\dim X = D(X; \mathbf{Q}(p))$, where $\mathbf{Q}(p)$ is the additive group of rationals mod 1 of the form m/p^s for any prime number p (V. Boltyanskiĭ, [5]); this result holds also when X is paracompact (Y. Kodama, *J. Math. Soc. Japan*, 18 (1966)).

G. Dimension and Measure

Let X be a separable metric space. Then $\dim X \leq n$ if and only if X is homeomorphic to a subset of a Euclidean space \mathbf{R}^{2n+1} whose $(n+1)$ -dimensional † Hausdorff measure is 0 (E. Szpilrajn, *Fund. Math.*, 28 (1937); also — [1, 2]). The infimum of $\alpha \geq 0$ such that the Hausdorff measure $\Lambda_{\alpha}(X)$ of dimension α vanishes is called the **Hausdorff dimension** of X .

H. Dimension Type (Fréchet's Definition)

In analogy to the theory of † cardinal numbers in set theory, M. Fréchet (1909) defined the dimension type of topological spaces as fol-

lows: Two spaces X and Y are said to have the same **dimension type** if X is homeomorphic to a subset of Y and Y is homeomorphic to a subset of X .

I. Infinite-Dimensional Spaces

If X is a metric space with $0 < \dim X < \infty$, then for each positive integer m with $m \leq \dim X$, X contains a (closed) subset S with $\dim S = m$. Tumarkin asked the following question: For an infinite-dimensional compact metric space X and for each positive integer m , does X contain a closed subset S with $\dim S = m$? D. Henderson (*Amer. J. Math.*, 89 (1967)) answered this question in the negative. Furthermore, J. Walsh (*Bull. Amer. Math. Soc.*, 84 (1978)) constructed an infinite-dimensional compact metric space X such that if S is an arbitrary subset of X with $\dim S > 0$ then $\dim S = \infty$.

References

- [1] K. Morita, Dimension theory (in Japanese), Iwanami, 1950.
- [2] W. Hurewicz and H. Wallman, Dimension theory, Princeton Univ. Press, 1941.
- [3] J. Nagata, Modern dimension theory, Noordhoff, second edition, 1965.
- [4] K. Nagami, Dimension theory, Academic Press, 1970.
- [5] P. S. Aleksandrov, The present status of the theory of dimension, *Amer. Math. Soc. Transl.*, (2) 1 (1955), 1–26. (Original in Russian, 1951.)
- [6] L. E. J. Brouwer, Beweis der Invarianz der Dimensionenzahl, *Math. Ann.*, 70 (1911), 161–165.
- [7] H. Lebesgue, Sur la non-applicabilité de deux domaines appartenant respectivement à des espaces à n et $n + p$ dimensions, *Math. Ann.*, 70 (1911), 166–168.
- [8] L. E. J. Brouwer, Beweis der Invarianz des n -dimensionalen Gebiets, *Math. Ann.*, 71 (1912), 305–313.
- [9] L. E. J. Brouwer, Über den natürlichen Dimensionsbegriff, *J. Reine Angew. Math.*, 142 (1913), 146–152.
- [10] P. Uryson, Les multiplicités cantorienes, *C. R. Acad. Sci. Paris*, 175 (1922), 440–442.
- [11] K. Menger, Dimensionstheorie, Teubner, 1928.
- [12] A. Pears, Dimension theory for general spaces, Cambridge Univ. Press, 1975.
- [13] R. Engelking, Dimension theory, North-Holland, 1978.
- [14] K. Morita, Dimension of general topological spaces, *Surveys in General Topology*, G. M. Reed (ed.), Academic Press, 1980.

118 (V.9) Diophantine Equations

A. General Remarks

A **Diophantine equation** is an [†]algebraic equation whose coefficients lie in the ring \mathbf{Z} of rational integers and whose solutions are sought in that ring. The name comes from Diophantus, an Alexandrian mathematician of the third century A.D., who proposed many Diophantine problems; but such equations have a very long history, extending back to ancient Egypt, Babylonia, and Greece. As early as the sixth century B.C., Pythagoras is said to have partially solved the equation $x^2 + y^2 = z^2$ by $x = 2n + 1$, $y = 2n^2 + 2n$, $z = y + 1$. A general solution is given by the Pythagorean numbers $x = m^2 - n^2$, $y = 2mn$, $z = m^2 + n^2$.

[†]Fermat's problem also concerns a Diophantine equation.

Systematic studies of Diophantine equations over \mathbf{Z} have been made for the linear equation $\sum_{i=1}^n a_i x_i = a$ ($a_i, a \in \mathbf{Z}$) and for the quadratic equation $ax^2 + bxy + cy^2 = k$ ($a, b, c, k \in \mathbf{Z}$) in two unknowns. The latter forms a principal topic of C. F. Gauss's *Disquisitiones arithmeticae* and can be regarded as a starting point of modern algebraic number theory. The special quadratic equation $t^2 - Du^2 = \pm 4$ ($D \in \mathbf{Z}$) is called **Pell's equation**. If $D < 0$, then Pell's equation has only a finite number of solutions. If $D > 0$, then all solutions t_n, u_n of Pell's equation are given by $\pm((t_1 + u_1\sqrt{D})/2)^n = (t_n + u_n\sqrt{D})/2$, provided that the pair t_1, u_1 is a solution with the smallest $t_1 + u_1\sqrt{D} > 1$ [15]. Using continued fractions (\rightarrow 83 Continued Fractions), we can determine t_1, u_1 explicitly. A general quadratic Diophantine equation $ax^2 + bxy + cy^2 = k$ with two unknowns can be solved completely if we use solutions of Pell's equation; this is an application of the arithmetic of quadratic fields (\rightarrow 347 Quadratic Fields) [1]. On quadratic Diophantine equations of several unknowns, there are deep studies by C. L. Siegel (\rightarrow 348 Quadratic Forms).

Diophantine problems consist of giving criteria for the existence of solutions of algebraic equations in rings and fields and eventually determining the number of such solutions. The fundamental ring of interest is \mathbf{Z} and the fundamental field of interest is \mathbf{Q} . One discovers rapidly, however, that to have all the technical maneuverability necessary for handling general problems, one must consider rings and fields of finite type over \mathbf{Z} and \mathbf{Q} . Furthermore, one is led to consider finite fields and local fields when one deals with a localization of the problems under consideration.

Techniques from various fields of mathematics, e.g., algebraic number theory, algebraic geometry, analysis, Diophantine approximation, etc., have been successfully applied to solve Diophantine problems. However, much remains unsolved today. Yu. V. Matiyasevich (1970) showed that Hilbert's tenth problem is unsolvable; there is no general method of telling whether a Diophantine equation has a solution. This theorem in a sense indicates the complexity of Diophantine problems. For many centuries, no other topic has engaged the attentions of so many mathematicians, both professional and amateur, or resulted in so many published papers. For these miscellaneous results — Dickson [1] and Mordell [2].

B. Equations over Finite Fields

Let k be a finite field of characteristic p consisting of q ($= p^f$) elements. **Chevalley's theorem:** Let f be a \dagger form of degree d in n variables with coefficients in k such that $d < n$. Then $f=0$ has a nontrivial solution in k . A generalization is **Warning's theorem:** Let f_1, \dots, f_t be polynomials with coefficients in k in n variables of degrees d_1, \dots, d_t , respectively, and suppose that $d = d_1 + \dots + d_t < n$. Then the number N of common zeros of f_1, \dots, f_t satisfies $N \equiv 0 \pmod{p}$. **Warning's second theorem** asserts that if $N > 0$ then $N \geq q^{n-d}$. Warning's theorem was also improved by J. Ax (1964) to the effect that $N \equiv 0 \pmod{q^b}$ for any integer $b < n/d$ [3]. For equations over finite fields, counting the number of solutions is important. Let $f(x, y)$ be an \dagger absolutely irreducible polynomial in x and y over k . Let N denote the number of zeros in k of $f(x, y)$. A. Weil proved $|N - q| \leq 2g\sqrt{q} + c(d)$, where g is the genus of the curve $f(x, y) = 0$ and $c(d)$ is a constant depending on d . Weil's proof requires the use of deep results from algebraic geometry. This theorem is equivalent to the \dagger Riemann hypothesis for algebraic curves over finite fields [4]. Later, using Stepanov's method, W. M. Schmidt and E. Bombieri (1973) independently gave new proofs which do not depend on algebraic geometry [3]. P. Deligne [5] proved a far-reaching generalization of Weil's theorem to \dagger nonsingular absolutely irreducible equations in n variables. He showed $|N - q^{n-1}| = O(q^{(n-1)/2})$. This is a part of the \dagger Weil conjecture for zeta functions of algebraic varieties over finite fields (\rightarrow Section E; 450 Zeta Functions Q). Schmidt obtained in an elementary manner a weaker estimate $|N - q^{n-1}| = O(q^{n-3/2})$ but without the assumption of nonsingularity (\rightarrow Section F).

C. Equations over Local Fields

A method of solving problems in number theory by use of embeddings of the ground field into its \dagger completions is called a local method. Such methods have important consequences when applied to Diophantine equations. Let f be a polynomial in n variables with rational integer coefficients. The congruence $f \equiv 0 \pmod{p^k}$ is solvable for all $k \geq 1$ if and only if $f = 0$ is solvable in $\dagger p$ -adic integers. This is an easy consequence of the compactness of the ring of p -adic integers. We can solve $f = 0$ in p -adic integers provided that we can solve an infinite sequence of congruences. It is generally difficult to tell when we may limit our consideration to only a finite number of these. In this respect, the following lemma is most useful. **Hensel's lemma:** Let $f(x_1, \dots, x_n)$ be a polynomial whose coefficients are p -adic integers. Let $\gamma_1, \dots, \gamma_n$ be p -adic integers such that for some i ($1 \leq i \leq n$) and an integer $\delta \geq 0$, we have $f(\gamma_1, \dots, \gamma_n) \equiv 0 \pmod{p^{2\delta+1}}$ and $\partial f / \partial x_i(\gamma_1, \dots, \gamma_n) \equiv 0 \pmod{p^\delta}$, $\not\equiv 0 \pmod{p^{\delta+1}}$. Then there exist p -adic integers $\theta_1, \dots, \theta_n$, such that $f(\theta_1, \dots, \theta_n) = 0$ and $\theta_i \equiv \gamma_i \pmod{p^{\delta+1}}$ for $i = 1, \dots, n$. The case $\delta = 0$ is often useful; it implies that a nonsingular solution mod p can be extended to a p -adic solution. Generalization to simultaneous equations is also known [6]. Skolem's method is sometimes useful when we investigate certain types of equations over \dagger local fields. This method is based on some simple properties of local analytic manifolds over local fields [7]. If a quadratic form has zeros in each local field, then it has a rational zero (\dagger Minkowski-Hasse theorem). When a theorem of this type holds, we say that the \dagger Hasse principle holds (\rightarrow 348 Quadratic Forms). For forms of higher degree, the Hasse principle no longer holds even if the forms are absolutely irreducible and nonsingular. Counterexamples were first found for cubic (E. S. Selmer, 1951) and quintic (M. Fujiwara, 1972) forms. Asymptotic formulas in Waring's problem (\rightarrow 4 Additive Number Theory E) can be regarded as an analytic form of the Hasse principle. As to the quantitative formulation of the Hasse principle, there are deep results of Siegel for quadratic forms and their generalization by T. Tamagawa. R. Brauer (1945) showed that forms in sufficiently many variables represent zero in all p -adic fields. Forms of odd degree represent zero in the field of rational numbers if the number of variables is sufficiently large compared with the degree (B. J. Birch, 1957). Let f be a polynomial with p -adic integer coefficients and c_m ($m \geq 0$) be the number of solutions to the congruence $f \equiv 0 \pmod{p^m}$. The series $\varphi(t) = \sum_{m=0}^{\infty} c_m t^m$ is called the Poincaré series of f .

Diophantine Equations

J. Igusa (1975) proved, by using his theory of asymptotic expansions together with Hiro-naka's †resolution theorem, that $\varphi(t)$ is a rational function of t [8].

D. Integral Solutions of Some Diophantine Equations

In this section we are concerned with those equations for which some "theory" exists. For isolated results \rightarrow [1, 2].

(1) Binary Forms. Thue's theorem (1908): If $f(x) = \sum_{v=0}^n a_v x^v$ ($a_v \in \mathbb{Z}, n > 2$) has distinct roots, then the number of rational integral solutions of $\sum_{v=0}^n a_v x^v y^{n-v} = a$ ($\mathbb{Z} \ni a \neq 0$) is finite. This theorem is a direct consequence of Thue's theorem on Diophantine approximation, which says that there are only a finite number of rational numbers p/q ($p, q \in \mathbb{Z}, q > 0$) with $|\alpha - p/q| < 1/q^{(n/2)+1}$ for a given algebraic number α of degree n ($n > 2$) [9, p. 122]. K. F. Roth proved that $(n/2) + 1$ in this formula can be replaced by $2 + \varepsilon$ (ε is an arbitrary positive number independent of n) (*Mathematika*, 2 (1955), 1–20). **Roth's theorem** was generalized to some cases of number fields and function fields (\rightarrow 182 Geometry of Numbers) and is applied to Diophantine equations [9, 10]. A. Baker (1968), using a completely different method, has given explicit upper bounds for the solutions of Thue's equations, thus enabling one to compute effectively all the solutions. More precisely, if f in Thue's equation is irreducible over the rationals, then every integer solution (x, y) of the equation satisfies $\max(|x|, |y|) < \exp((nH)^{(10n)^5} + (\log a)^{2n+2})$, where H is the †height of f . The proof of this remarkable theorem is based on his deep result concerning the lower bound for the linear forms in the logarithm of algebraic numbers (*Mathematica*, 13 (1966); 14 (1967)). Baker's method has been applied to elliptic, hyperelliptic, and other curves (Baker, H. Stark, J. Coates, V. G. Sprindzhuk, etc.).

(2) Higher-Degree Forms. A natural generalization of binary forms is a **norm form**. Let K be an algebraic number field of degree $t \geq 3$ and $\alpha_1, \dots, \alpha_n$ be elements of K . Then the norm $N(\alpha_1 x_1 + \dots + \alpha_n x_n) = \prod_{i=1}^t (\alpha_1^{(i)} x_1 + \dots + \alpha_n^{(i)} x_n)$, where $\alpha^{(i)}$ denotes a conjugate of α , is a form of degree t with rational coefficients. It is easy to see that every form which has rational coefficients and is irreducible over \mathbb{Q} but which is a product of linear forms with algebraic coefficients is a constant multiple of a norm form [7]. A module M in an algebraic number field K is called **degenerate** if M has a submodule N such that, for some $\alpha \in K$, αN is a full module

in some subfield K' of K , where K' is neither \mathbb{Q} nor an imaginary quadratic field. The most basic in the theory of norm forms is (W. M.) **Schmidt's theorem**: Let $\alpha_1, \dots, \alpha_n$ be linearly independent over \mathbb{Q} and suppose that the module generated by $\alpha_1, \dots, \alpha_n$ is nondegenerate; then $N(\alpha_1 x_1 + \dots + \alpha_n x_n) = c$ ($c \in \mathbb{Q}$) has only finitely many solutions in integers x_1, \dots, x_n . (*Math. Ann.*, 191 (1971)). The proof is based on his remarkable result on †simultaneous approximation which generalizes Roth's theorem (\rightarrow 182 Geometry of Numbers G). There are investigations on special norm forms (T. Skolem, N. I. Fel'dman, K. Ramanathan, K. Györy, M. Fujiwara, etc.). For general forms of higher degree, not much is yet known except for the additive forms (\rightarrow 4 Additive Number Theory E). H. Davenport [11] proved that if $f(x)$ is a cubic form with rational integer coefficients in n variables, then $f(x) = 0$ has a nontrivial integral solution, provided that $n \geq 16$. This theorem was proved by means of an exquisite application of the †circle method together with some geometry of numbers. A well-known conjecture is that $n \geq 10$ instead of $n \geq 16$. It is known that over local fields, any cubic form in 10 variables has a nontrivial zero. There are various results of this type for simultaneous additive, quadratic, and cubic forms (Davenport, D. Lewis, R. Cook, Schmidt, etc.) [17]. A satisfactory theory of forms of higher degree, like that of quadratic forms, is not yet known but is quite desirable. In this vein, Igusa has obtained some new results of considerable interest, e.g., a †Poisson summation formula for higher-degree forms, using his theory of asymptotic expansions [8].

(3) Algebraic Curves. The fundamental result is **Siegel's theorem** (1929): Assume that the equations $f_i(X_1, \dots, X_n) = 0$ ($1 \leq i \leq m$) determine an algebraic curve with a positive †genus in an †affine space of dimension n . Then the number of rational integral solutions of $f_i(X_1, \dots, X_n) = 0$ ($1 \leq i \leq m$) is finite. This theorem was generalized by S. Lang in the following form: Let K be a finitely generated field over \mathbb{Q} and I a subring of K that is finitely generated over \mathbb{Z} . Furthermore, let C be a nonsingular projective algebraic curve with a positive genus defined over K , and let φ be a rational function on C defined over K . Then there are only a finite number of points P on C with $\varphi(P) \in I$ [10]. The proof of this theorem is based on a generalization of Roth's theorem in the above sense and on the weak Mordell-Weil theorem (\rightarrow Section C). A. Robinson and P. Roquette gave another approach to Siegel's theorem from the standpoint of nonstandard arithmetic (*J. Number Theory* 7 (1975)). On the other hand, a necessary condition for the exis-

tence of infinitely many solutions of $f(X, Y) = 0$ with rational integral coefficients was given by C. Runge (*J. Reine Angew. Math.*, 100 (1887)).

(4) Elliptic Curves. An elliptic curve E is an \dagger Abelian variety of dimension 1, or what is the same, an irreducible nonsingular \dagger projective algebraic curve of \dagger genus 1 furnished with a point O as origin. The \dagger Riemann-Roch theorem defines a group law on the set of \dagger divisor classes of E . Actually, if P, P' are points of E , then there exists a unique point P'' such that $(P) + (P') \sim (P'') + (O)$, where \sim means linear equivalence, i.e., the left-hand side minus the right-hand side is the divisor of a rational function on the curve. The group law on E is then defined by $P + P' = P''$. If the characteristic $\neq 2$ or 3, using the Riemann-Roch theorem one finds that the curve E can be defined by a Weierstrass equation $y^2 = x^3 + ax + b$ with a, b in the ground field over which the curve is defined. Conversely, any homogeneous nonsingular cubic equation has genus 1 and defines an elliptic curve in the projective plane once the origin has been selected. If both the curve and the origin are defined over a field k , then the group law is also defined over k , and it becomes a 1-dimensional Abelian variety defined over k . If the ground field k is the field of complex numbers, the group law is the same as that given by the \dagger addition theorem of the \dagger Weierstrass \wp -function with invariants $g_2 = -4a$ and $g_3 = -4b$ through the parametrization $x = \wp(u)$, $y = \frac{1}{2}\wp'(u)$. Much of the Diophantine theorems on elliptic curves are generalized to Abelian varieties. Here we shall deal mainly with elliptic curves defined by Weierstrass equations over \mathbf{Q} . Extension to algebraic number fields usually causes no trouble. For more general elliptic curves \rightarrow [19]. The Lutz-Mattuck theorem (\rightarrow Section E) obviously implies that the points of finite order in E_k ($k = \mathbf{Q}_p$) form a finite group. This torsion group is computable. In case a and b are in \mathbf{Z} then any point of finite order in E_0 has coordinates (x, y) in \mathbf{Z} and, if $y \neq 0$, $y^2 \mid 4a^3 + 27b^2$ (Lutz-Nagell). The **WC group** (Weil-Châtelet group) of E over k is the birational class of principal homogeneous spaces over k . The extent of validity of the Hasse principle for elliptic curves can be measured by the **Tate-Shafarevich group**, which is defined as the set of elements of the WC group that are everywhere locally trivial. This group is conjectured to be a finite group. For other results and interesting conjectures \rightarrow [12–14]. The number of integral points on elliptic curves is finite according to Siegel's theorem on algebraic curves. Explicit bounds for the size of these points have been given for several types of

elliptic curves by using Baker's method. For example, if $f(x, y)$ is an absolutely irreducible polynomial with coefficients in \mathbf{Z} such that the curve $f = 0$ has genus 1, then $\max(|x|, |y|) < \exp \exp \exp((2H)^m)$, where $m = 10^{d^{10}}$, $d = \deg f$, and H is the height of f (Baker and Coates, 1970. The method of proof was to reduce it to the Weierstrass equation case, which had been treated earlier by Baker, with a better bound.) By the \dagger Mordell-Weil theorem (\rightarrow Section E), $A_{\mathbf{Q}} \cong \mathbf{Z}^r \times$ finite torsion group. Here r is called the **rank** of E over \mathbf{Q} . There is a rather doubtful conjecture to the effect that the rank r is bounded. The rank r is conjectured to be equal to the order of the zero of $L(s, E)$ at $s = 1$ (**Birch-Swinnerton-Dyer conjecture**). Much numerical and theoretical evidence supports this famous conjecture [13].

E. Rational Points of Algebraic Varieties

Let V be an \dagger abstract algebraic variety defined over a field k , and let P be a point of V . Then P is called a **rational point** over k of V if the coordinates of the \dagger representative P_α contained in an \dagger affine open set V_α of V are in k (\rightarrow 16 Algebraic Varieties D). This definition is independent of the choice of the representative P_α . In particular, if V is a \dagger projective variety, the point P given by the \dagger homogeneous coordinates (x_0, x_1, \dots, x_n) is rational if and only if $x_i/x_p \in k$ ($0 \leq i \leq n$, $x_p \neq 0$). In the following we state main results concerning rational points of algebraic varieties, especially results concerning \dagger Abelian varieties, restricting k to be either an \dagger algebraic number field of finite degree, a \dagger p-adic number field, or a \dagger finite field.

Mordell-Weil Theorem. Let A be an Abelian variety of dimension n defined over an algebraic number field k of finite degree. Then the group A_k of all k -rational points on A is finitely generated. This theorem was proved by L. J. Mordell (1922) for the case of $n = 1$ and by Weil (1928) for the general case [10]. The assertion that A_k/mA_k is a finite group for any rational integer m is called the **weak Mordell-Weil theorem**; this theorem is basic in the proof of the Mordell-Weil theorem and is used in the proof of Siegel's theorem, too. A generalization of the Mordell-Weil theorem is obtained when k is a field (of arbitrary characteristic) finitely generated over the \dagger prime field [10].

If A is defined over a finite algebraic number field k , we have the following conjectures of Birch, Swinnerton-Dyer, and Tate on the rank of A_k . Let \mathfrak{p} be a prime ideal of k at which A has a good \dagger reduction, and denote by $A_{\mathfrak{p}}$ the reduced variety. Let $\pi_{\mathfrak{p}}^{(1)}, \dots, \pi_{\mathfrak{p}}^{(2n)}$ be the eigen-

values of the $N(p)$ th power endomorphism of A_p with respect to an l -adic representation, where $N(p)$ denotes the norm of p (\rightarrow 3 Abelian Varieties E, N), and put $L_p(s, A) = \prod_{i=1}^{2n} (1 - \pi_p^{(i)} N(p)^{-s})^{-1}$. The L -function of A defined by $L(s, A) = \prod' L_p(s, A)$, where the product ranges over all good primes, is the principal part of the zeta function of A (\rightarrow 450 Zeta Functions S). Birch and Swinnerton-Dyer conjectured that if $k = \mathbf{Q}$ and A is of dimension 1, then there exists a constant $C \neq 0$ such that $L(s, A) \sim C(s-1)^g$ as $s \rightarrow 1$. Tate generalized this conjecture to any A and k . Moreover, the constant C , appropriately modified by factors corresponding to the bad primes and the infinite primes, is thought to be expressible in terms of certain arithmetic invariants of A [13].

Lutz-Mattuck Theorem. The group of rational points of an Abelian variety A of dimension n over a p -adic number field k contains a subgroup of finite index isomorphic to the direct sum of n copies of the p -adic integers in k (E. Lutz, *J. Reine Angew. Math.*, 177 (1937); A. Mattuck, *Ann. Math.*, 62 (1955)).

Mordell's Conjecture. In his 1922 paper, in which the above theorem on the set of rational fields on 1-dimensional Abelian varieties (i.e., on elliptic curves) was established, Mordell stated the conjecture: Any algebraic curve of genus $g \geq 2$ defined over \mathbf{Q} has only a finite number of rational points. The same can be conjectured for such curves defined over any algebraic number field k of finite degree. This had remained as a conjecture until 1983. In 1961 I. R. Shafarevich conjectured: Let k be any algebraic number field of finite degree, S a finite set of finite prime spots of k , and g any natural number ≥ 2 . Then there are, up to k -isomorphism, only a finite number of non-singular algebraic curves of genus g defined over k having good reduction at every finite prime spot outside S .

In 1973, A. N. Parshin showed that Mordell's conjecture followed from this conjecture, which was finally proved in 1983 by G. Faltings [7]. Mordell's conjecture was thus settled in the affirmative. Analogs of these conjectures on algebraic function fields over finite fields are easier than the original ones and had been proved in the 1960s for Mordell's conjecture by Yu. I. Manin, H. Grauert, and M. Miwa, and for Shafarevich's conjecture by S. Alakelov, A. N. Parshin, and L. Szpiro.

F. C_i -Fields

Let F be a field, and let $i \geq 0$, $d \geq 1$ be integers. Let f be a homogeneous polynomial of n variables of degree d with coefficients

in F . If the equation $f=0$ has a solution $(x_1, \dots, x_n) \neq (0, \dots, 0)$ in F for any f with $n > d^i$, then F is called a $C_i(d)$ -field. If F is a $C_i(d)$ -field for any $d \geq 1$, then F is called a C_i -field. In order for F to be a C_0 -field, it is necessary and sufficient that F be an $\bar{\mathbf{A}}$ -algebraically closed field. A C_1 -field is sometimes called a **quasi-algebraically closed field**. There exists no non-commutative algebra over a C_1 -field F . A finite field is C_1 (C. Chevalley (1936)). If F_0 is algebraically closed, then $F = F_0(X)$ (rational function field of one variable) is a C_1 -field (**Tsen's theorem**). A homogeneous polynomial f of $n = d^i$ variables of degree d with coefficients in F such that $f=0$ has no solution in F except $(0, \dots, 0)$ is called a **normic form** of order i in F . If a C_i -field F_0 has at least one normic form of order i , then (i) $F_0(X_1, \dots, X_k)$ is a C_{i+k} -field; and (ii) an extension of F_0 of finite degree is a C_i -field. A complete field F with respect to an $\bar{\mathbf{A}}$ -exponential valuation is a C_1 -field whenever its residue field F_0 is algebraically closed. The field F of power series of one variable over a finite field F_0 is a C_2 -field (Lang). E. Artin conjectured that a p -adic field \mathbf{Q}_p is a C_2 -field. It was proved by H. Hasse (1923) that \mathbf{Q}_p is a $C_2(2)$ -field and by D. Lewis (1952) that \mathbf{Q}_p is a $C_2(3)$ -field. However, G. Terjanian (1966) [17] gave a counterexample to Artin's conjecture; that is, he gave a quartic form of 18 variables with coefficients in \mathbf{Q}_2 having only trivial zero in \mathbf{Q}_2 . Ax and S. Kochen (1965) [18] proved that for any integer $d \geq 1$ there exists an integer $p_0(d)$ such that \mathbf{Q}_p is a $C_2(d)$ -field for $p > p_0(d)$ (\rightarrow 276 Model Theory E).

References

- [1] L. Dickson, History of the theory of numbers I–III, Chelsea, 1952.
- [2] L. J. Mordell, Diophantine equations, Academic Press, 1969.
- [3] W. M. Schmidt, Equations over finite fields, Lecture notes in math. 536, Springer, 1976.
- [4] A. Weil, Sur les courbes algébriques et les variétés qui s'en déduisent, Actualités Sci. Ind., 1041 (1948).
- [5] G. Faltings, Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, Inventiones Math., 73 (1983), 349–366.
- [6] M. Greenberg, Lectures on forms in many variables, Benjamin, 1969.
- [7] Z. I. Borevich and I. R. Shafarevich, Number theory, Academic Press, 1966. (Original in Russian, 1964.)
- [8] J. Igusa, Lectures on forms of higher degree, Tata Inst., 1978.
- [9] W. J. LeVeque, Topics in number theory, Addison-Wesley, 1956.

- [10] S. Lang, Diophantine geometry, Interscience, 1962.
- [11] H. Davenport, Analytic methods for Diophantine equations and Diophantine inequalities, Univ. of Michigan, 1962.
- [12] J. W. S. Cassels, Diophantine equations with special reference to elliptic curves, J. London Math. Soc., 41 (1966), 193–291.
- [13] H. P. F. Swinnerton-Dyer, The conjectures of Birch and Swinnerton-Dyer and of Tate, Proc. Conf. on Local Fields, Driebergen, Springer, 1967.
- [14] S. Lang, Elliptic curves: Diophantine analysis, Springer, 1978.
- [15] I. Niven and H. Zuckerman, An introduction to the theory of numbers, Wiley, 1960.
- [16] J. Joly, Equations et variétés algébriques sur un corps fini, Enseignement Math., (2) 19 (1973), 1–117.
- [17] D. Lewis, Diophantine equations: p -adic method, Studies in Math., vol. 6, Math. Assoc. Amer., 1969.
- [18] J. Ax and S. Kochen, Diophantine problems over local fields I, II, Amer. J. Math., 87 (1965), 605–630, 631–648.
- [19] J. Tate, The arithmetic of elliptic curves, Inventiones Math., 23 (1974), 170–206.

119 (XXI.20) Dirichlet, Peter Gustav Lejeune

Peter Gustav Lejeune Dirichlet (February 13, 1805–May 5, 1859) was born of a French family in Düren, Germany. From 1822 to 1827 he was in Paris, where he became a friend of J.-B. Fourier. In 1827, he was appointed lecturer at the University of Breslau; in 1829, lecturer at the University of Berlin; and in 1839, professor at the University of Berlin. In 1855, he was invited to succeed C. F. Gauss at the University of Göttingen, where he spent his last four years as a professor.

His works cover many aspects of mathematics; however, those on number theory, analysis, and potential theory are the most famous. He greatly admired Gauss and is said to have kept Gauss's *Disquisitiones arithmeticae* at his side even when traveling.

In number theory, he created the Dirichlet series and proved that a sequence in arithmetic progression contains infinitely many prime numbers, provided that the first term and the common difference are relatively prime. Also, using his “drawer principle,” which states that if there are $n + 1$ objects in n drawers then at least one drawer contains at least 2 objects, he clarified the structure of unit groups of algebraic

120 A Dirichlet Problem

braic number fields. In potential theory he dealt with the Dirichlet problem concerning the existence of harmonic functions. He also gave Dirichlet's condition for the convergence of trigonometric series.

References

- [1] P. G. L. Dirichlet's Werke I, II, Reimer, 1889, 1897 (Chelsea, 1969).
- [2] P. G. L. Dirichlet, Vorlesungen über Zahlentheorie, with supplements by R. Dedekind, Vieweg, fourth edition, 1894 (Chelsea, 1968).
- [3] F. Klein, Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert I, Springer, 1926 (Chelsea, 1956).

120 (X.30) Dirichlet Problem

A. The Classical Dirichlet Problem

Let D be a bounded or unbounded domain in \mathbf{R}^n ($n \geq 2$) with compact boundary S . The classical **Dirichlet problem** is the problem of finding a harmonic function in D that assumes the values of a prescribed continuous function on S . This problem is also called the first boundary value problem (\rightarrow 193 Harmonic Functions and Subharmonic Functions). In this article f always stands for a boundary function given on S . The problem is called an **interior problem** if D is bounded and an **exterior problem** if D is unbounded. In an exterior problem, it is further required that when an inversion with center at an exterior point P_0 of D is performed on D and a Kelvin transformation is performed on the solution in D (when the solution exists), the function thus obtained on the inverted image of D be harmonic at P_0 ($n \geq 3$). When $n = 2$, the solution in D , which is already regarded as a function on the inverted image of D , is required to be harmonic at P_0 . Thus an interior problem can be transformed to an exterior problem, and vice versa. We now explain the history of the classical problem.

Let D be a bounded domain with boundary S in \mathbf{R}^3 . G. Green (1828) asserted that if S is sufficiently smooth,

$$u(P) = -\frac{1}{4\pi} \int_S f(Q) \frac{\partial G(P, Q)}{\partial n_Q} d\sigma(Q) \quad (1)$$

is the solution for the Dirichlet problem, where f is prescribed on S , $G(P, Q)$ is Green's function with the pole at Q in D , n_Q is the outward-drawn normal to S at Q , and $d\sigma$ is the surface

face element on S . He took for granted the existence of Green's function from physical consideration of the problem. Thus his discussion was not quite rigorous. This defect was corrected by A. M. Lyapunov (1898) under a certain condition on S . Denote by u_m the Newtonian potential of a measure with density $m \geq 0$ on S . Assume that a continuous function f on S and a positive constant a are given. In 1840, C. F. Gauss investigated the existence of a density $m_f \geq 0$ on S of total mass a which satisfies $\int_S (u_{m_f} - 2f)m_f d\sigma = \min_m \int_S (u_m - 2f)m d\sigma$, where the total mass of m is equal to a . He asserted also that $u_{m_f} - f$ is equal to a constant b on S . If $f \equiv 0$, then u_{m_0} must be equal to a positive constant c on S , and hence $u_{m_f} - bc^{-1}u_{m_0}$ must be a solution of the exterior problem for the boundary function f on S . However, his discussion was incomplete because we cannot always ensure the existence of a density that gives a measure minimizing the integral. Moreover, even in the case where D is a ball, there exists a continuous function $f \geq 0$ on S such that there is no Newtonian potential that is equal to f on S up to a constant (M. Ohtsuka, 1961). Gauss (1840), W. Thomson (Lord Kelvin) (1847), and G. L. Dirichlet solved the Dirichlet problem by making use of the so-called **Dirichlet principle**, which is explained in detail in Section F. After K. Weierstrass (1870) pointed out that there is a case where no minimizing function exists, D. Hilbert (1899) gave a rigorous proof of the Dirichlet principle under a certain condition. Meanwhile, C. G. Neumann (1870) solved the Dirichlet problem rigorously for the first time, although he assumed that D is a convex domain with a smooth boundary. First, he considered the potential $W_1 = (1/2\pi) \cdot \int_S f(\partial r^{-1}/\partial n) d\sigma$ of a double layer in D ; then he formed the potential $W_2 = (1/2\pi) \cdot \int_S f_1(\partial r^{-1}/\partial n) d\sigma$ of a double layer with the values f_1 of W_1 on S and defined W_3, W_4, \dots similarly. The series $W_1 - W_2 + W_3 - W_4 + \dots$ plus a suitable constant gives a solution to the exterior Dirichlet problem for the boundary function f . In 1887, H. Poincaré also used (1) to solve the Dirichlet problem. He obtained Green's function with the pole at O in a bounded domain D in the following manner: Let D' be the image of D by an inversion with center O , S_0 be a spherical surface surrounding the boundary $\partial D'$ of D' , and μ be a uniform measure on S_0 such that the potential of μ is equal to 1 inside S_0 . By "sweeping out μ to $\partial D'$ ", the solution in D' of the exterior problem for the boundary function 1 is obtained. A ^{*}Kelvin transformation of this solution yields the solution h in D of the interior problem for the boundary function $1/OP$. Now $1/OP - h(P)$ is Green's function in D . In 1899

Poincaré used another method (without utilizing (1)) to solve the Dirichlet problem [8]. He observed that it is sufficient to consider the case where f is equal to the restriction to S of a polynomial g and that g is expressed in D as the sum of the Newtonian potential of a signed measure τ with density $-\Delta g/(4\pi)$ and a function that is harmonic in D and continuous on $D \cup S$. If it is possible to sweep out τ to ∂D , then the solution is obtained. He showed that this is in fact the case if at every point P of S there exists a cone that is disjoint from D and has its vertex at P . This condition is called **Poincaré's condition**. In 1900, I. Fredholm discussed the Dirichlet problem by reducing it to a problem of integral equations. A domain D is called a **Dirichlet domain** if the (classical) Dirichlet problem is always solvable in D . H. Lebesgue (1912) showed that a solution is obtained by the method of iterative averaging in every Dirichlet domain.

B. The Dirichlet Problem in a General Domain

It has been believed that the classical Dirichlet problem is always solvable in every domain until S. Zaremba observed in 1909 that the problem is not always solvable for a punctured ball. In 1913, Lebesgue gave a decisive example in which the domain is homeomorphic to a ball and bounded by a surface sufficiently smooth except at one point. Thus the central interest shifted to finding a harmonic function in D that depends only on a continuous function f given on ∂D and coincides with the classical solution when D is a Dirichlet domain. Extend f to a continuous function in the whole space, and denote it by f_0 . Approximate D by an increasing sequence $\{D_n\}$ of Dirichlet domains, and denote by u_n the solution in D_n of the Dirichlet problem for the boundary function f_0 . N. Wiener proved in 1924 that u_n converges to a harmonic function that is independent of the choice of the extension of f and $\{D_n\}$. The problem of deciding where on ∂D the general solution assumes the given boundary values is treated in Section D. O. D. Kellogg (1928) found a general method that includes Poincaré's method of sweeping out, Schwarz's alternating method, and the result of Wiener. Both Poincaré's method of sweeping out and Lebesgue's method of iterative averaging yield Wiener's general solution.

C. Perron's Method

We explain O. Perron's method (1923) by considering the improved method of M. Brelot [1]. For simplicity we assume that the domain

D is bounded in \mathbf{R}^3 . Let U be the family of subharmonic functions u bounded above and satisfying $\limsup_{P \rightarrow M} u(P) \leq f(M)$ for any boundary point M . Define $H_f(P)$ as $\sup_u u(P)$, where u runs through U , if this family is not empty; otherwise, set $H_f \equiv -\infty$. Call H_f a **hypofunction**. Define \bar{H}_f by $-\bar{H}_f$ and call it a **hyperfunction**. If $H_f = \bar{H}_f$, the common function is denoted by H_f ; if $H_f(P) < \infty$, then H_f is harmonic. This is called a **Perron-Brelot solution** (**Perron-Wiener-Brelot solution** or simply **PWB solution**). The method of defining H_f is called **Perron's method** (the Perron-Brelot method or the Perron-Wiener-Brelot method).

Wiener showed in 1923 that the †Daniell-Stone integral can be regarded as a general solution if a †Daniell-Stone integrable function f is given on the boundary of a Dirichlet domain; in 1925, he showed that the same is true for a general domain (not necessarily Dirichlet). He showed also that his solution coincides with the Perron-Brelot solution H_f if f is continuous. Unfortunately, however, from a wrong example he concluded that $H_f \neq \bar{H}_f$ can hold even for a simple discontinuous f , and so he lost interest in Perron's method. Brelot (1939) corrected Wiener's erroneous conclusion and proved that the Daniell upper and lower integrals are equal to \bar{H}_f and H_f , respectively. To any continuous f there corresponds an H_f , and there exists a †Radon measure μ_P satisfying $H_f(P) = \int f d\mu_P$. This measure is called a **harmonic measure** or harmonic measure function. Brelot showed that $H_f = \bar{H}_f$ if and only if f is μ_P -integrable for one (or every) P . In particular, if D is a Dirichlet domain and E is a closed set on the boundary ∂D , then the harmonic measure function $\mu_P(E)$ takes the value 1 at an inner point (in the space ∂D) of E and vanishes on $\partial D - E$. We note that μ_P is equal to the measure obtained by sweeping out the unit mass at P to ∂D .

D. Regular Boundary Points

If $H_f(P) \rightarrow f(M)$ as $P \rightarrow M \in \partial D$ for any continuous function f on ∂D , then M is called **regular**. The regularity of a point is a local property. A boundary point that is not regular is called **irregular**. The regularity of M is equivalent to the convergence of μ_P as $P \rightarrow M$ to the unit mass at M with respect to the †vague topology. There are many sufficient conditions and necessary conditions for a boundary point to be regular. The existence of a **barrier** is a qualitative condition that is necessary and sufficient for a boundary point to be regular. It was used by Poincaré and so named and used effectively by Lebesgue. A barrier is a continuous superharmonic func-

tion in D that assumes the boundary value 0 at M and has a positive lower bound outside every ball with center at M . A positive superharmonic function defined in the intersection of D and a neighborhood of M and taking the boundary value 0 can be used as a barrier. A necessary and sufficient condition for a boundary point M to be regular is the existence of a Green's function in D assuming the value 0 at M . This condition was given by G. Bouligand (1925), and it follows from the existence of a barrier. Another necessary and sufficient condition of a quantitative nature was obtained by Wiener. It is equivalent to the requirement that the complement of D is not †thin at M (\rightarrow 338 Potential Theory G). Kellogg conjectured that the set of irregular boundary points is of capacity zero and verified this in \mathbf{R}^2 (1928). The conjecture was proved first by G. C. Evans (1933) in \mathbf{R}^3 , and different proofs were given by F. Vasilescu (1935) and O. Frostman (1935). The conjecture is also true in \mathbf{R}^n for $n \geq 4$.

E. The More General Dirichlet Problem

So far, we have been concerned with \mathbf{R}^n . More generally, Brelot and G. Choquet [3] obtained the following result in a Green space \mathcal{E} (\rightarrow 193 Harmonic Functions and Subharmonic Functions): Consider a metric space that contains \mathcal{E} and in which \mathcal{E} is everywhere dense, and denote by Δ the complement of \mathcal{E} with respect to the space. Let $\{F\}$ be a family of †filters on \mathcal{E} such that each F converges to a certain point of Δ . Suppose that $u \leq 0$ whenever u is subharmonic and bounded above on \mathcal{E} and $\limsup u \leq 0$ along every F . Assume the existence of a barrier v in a neighborhood in \mathcal{E} of the limit point Q of every F ; that is, v is to be positive superharmonic, to tend to 0 along F , and to have a positive lower bound outside every neighborhood of Q . Under these assumptions, we obtain the PWB solution on \mathcal{E} as in \mathbf{R}^3 . There are various examples of Δ and F that satisfy these conditions. In particular, L. Naïm [6] investigated in detail the case where Δ is the †Martin boundary. More generally, it is possible to treat the Dirichlet problem axiomatically (\rightarrow 193 Harmonic Functions and Subharmonic Functions).

F. The Dirichlet Principle

Let D be a bounded domain with a sufficiently smooth boundary in \mathbf{R}^n and f be a piecewise C^1 -function on D with finite **Dirichlet integral** $\|f\|^2 = \int_D |\text{grad } f|^2 d\tau$, where $d\tau$ is the volume element. Suppose that f has a continuous boundary value ϕ on ∂D . The classical Dirich-

let principle asserts that the solution of the Dirichlet problem for φ has the smallest Dirichlet integral among the functions that are piecewise of class C^1 in D and assume the boundary value φ . In a general domain, H_φ minimizes $\|u - f\|$ among harmonic functions u in D . Brelot [2] discussed the principle for a family of competing functions that are defined in a domain in \mathcal{E} and whose boundary values cannot be defined in the classical manner.

References

- [1] M. Brelot, Familles de Perron et problème de Dirichlet, Acta Sci. Math. Szeged., 9 (1939), 133–153.
- [2] M. Brelot, Etude et extensions du principe de Dirichlet, Ann. Inst. Fourier, 5 (1955), 371–419.
- [3] M. Brelot and G. Choquet, Espaces et lignes de Green, Ann. Inst. Fourier, 3 (1952), 199–263.
- [4] R. Courant, Dirichlet's principle, conformal mapping, and minimal surfaces, Interscience, 1950.
- [5] O. D. Kellogg, Foundations of potential theory, Springer, 1929.
- [6] L. Naïm, Sur le rôle de la frontière de R. S. Martin dans la théorie du potentiel, Ann. Inst. Fourier, 7 (1957), 183–281.
- [7] M. Ohtsuka, Dirichlet problem, extremal length and prime ends, Lecture notes, Washington Univ., St. Louis, 1962–1963.
- [8] H. Poincaré, Théorie du potentiel Newtonien, Leçons professées à la Sorbonne, Gauthier-Villars, 1899.
- [9] L. L. Helms, Introduction to potential theory, Interscience, 1969.

121 (XI.3) Dirichlet Series

A. Dirichlet Series

For $z = x + iy$, $\lambda_n > 0$, and $\lambda_n \uparrow +\infty$, the series of the form

$$f(z) = \sum_{n=1}^{\infty} a_n \exp(-\lambda_n z) \quad (1)$$

is called a **Dirichlet series** (more precisely, a **Dirichlet series of the type** $\{\lambda_n\}$). If $\lambda_n = n$, then (1) is a power series with respect to e^{-z} . If $\lambda_n = \log n$, the series (1) becomes

$$\sum_{n=1}^{\infty} a_n/n^z, \quad (2)$$

which is called an **ordinary Dirichlet series**. If $a_n = 1$, then (2) is the \dagger Riemann zeta function.

Series of the form (2) were introduced by P. G. L. Dirichlet in 1839 and utilized in an investigation of the problems of analytic number theory. Later J. Jensen (1884) and E. Cohen (1894) extended the variable z to complex numbers. The Dirichlet series is not only a useful tool in analytic number theory, but is also investigated as a generalization of power series. The \dagger Laplace transform is the generalization of the Dirichlet series to the integral, and similar formulas often hold for both cases.

B. Convergence Regions

If the series (1) converges at $z = z_0$, then it converges in the half-plane $\operatorname{Re} z > \operatorname{Re} z_0$. Therefore there is a uniquely determined real number S such that (1) converges in $\operatorname{Re} z > S$ and diverges in $\operatorname{Re} z < S$. If (1) always converges (diverges), we put $S = -\infty$ ($+\infty$). We call S the **abscissa of convergence** (or **abscissa of simple convergence**). Similarly, there is a uniquely determined real number A such that (1) converges absolutely in $\operatorname{Re} z > A$ and is not absolutely convergent in $\operatorname{Re} z < A$. We call A the **abscissa of absolute convergence**. Furthermore, there is a uniquely determined real number U such that (1) converges uniformly in $\operatorname{Re} z > U'$ for every $U' > U$ and does not converge uniformly in $\operatorname{Re} z > U''$ for every $U'' < U$. The number U is called the **abscissa of uniform convergence**. Among these abscissas we always have the relations

$$-\infty \leq S \leq U \leq A + \infty,$$

$$A - S \leq \limsup_{n \rightarrow \infty} \frac{\log n}{\lambda_n}.$$

The latter was proved by Cohen (1894). The numbers S, A, U are determined from a_n, λ_n by means of the formulas

$$S = \limsup_{x \rightarrow \infty} \frac{1}{x} \log \left| \sum_{[x] \leq \lambda_n < x} a_n \right|, \quad (3)$$

$$A = \limsup_{x \rightarrow \infty} \frac{1}{x} \log \left(\sum_{[x] \leq \lambda_n < x} |a_n| \right), \quad (4)$$

$$U = \limsup_{x \rightarrow \infty} \frac{1}{x} \log T_x,$$

$$T_x = \sup_{-\infty < y < +\infty} \left| \sum_{[x] \leq \lambda_n < x} a_n \exp(-i\lambda_n y) \right|, \quad (5)$$

where $[]$ is the \dagger Gauss symbol. Formulas (3) and (4) were proved by T. Kojima (1914) and (5) by M. Kunieda (1916). In particular, when $\lim_{n \rightarrow \infty} (\log n)/\lambda_n = 0$, we have

$$S = U = A = \limsup_{n \rightarrow \infty} \frac{\log |a_n|}{\lambda_n} \quad (6)$$

(O. Szász, 1922; G. Valiron, 1924).

The series (1) converges uniformly in the angular domain $\{z \mid |\arg(z - z_0)| < \alpha < \pi/2\}$, where the vertex z_0 lies on the line $\operatorname{Re} z_0 = S$. Hence it represents a holomorphic function in the domain $\operatorname{Re} z > S$, but it is possible that there is no singularity on the line $\operatorname{Re} z = S$. For example, if $a_n = (-1)^n$, then the series (2) has $S = 0$, but the sum is an entire function $(2^{1-z} - 1)\zeta(z)$. Taking the analytic continuation $f(z)$ of the series (1), the infimum R of ρ such that $f(z)$ is holomorphic in $\operatorname{Re} z > \rho$ is called the **abscissa of regularity**. It is still possible that there is no singularity on the line $\operatorname{Re} z = R$. We always have $R \leq S$, and R is given by

$$R = \sup_{-\infty < y < \infty} \lim_{x \rightarrow -\infty} (\log \log^+ |\varphi(x + iy)| + x),$$

$$\varphi(z) = \sum_{n=1}^{\infty} \frac{a_n \exp(-\lambda_n z)}{\Gamma(1 + \lambda_n)}, \quad (7)$$

where $\log^+ a = \max(\log a, 0)$ (C. Tanaka, 1951). The infimum B of ρ such that $f(z)$ is bounded in $\operatorname{Re} z > \rho$ is called the **abscissa of boundedness**. We always have $R \leq B \leq A$. H. Bohr proved the following three theorems concerning these values: (1) If $\{\lambda_n\}$ are linearly independent over the ring of integers, then $A = B$ (1911). (2) If $(\lambda_{n+1} - \lambda_n)^{-1} = O(\exp \varepsilon^{\lambda_n})$ for every $\varepsilon > 0$, then $U = B$ (1913). (3) If $\limsup (\log n)/\lambda_n = 0$, then $S = U = A = B$ (1913). In the final case, the values are given by (6).

C. Properties of Functions Given by Dirichlet Series

The coefficients a_n in (1) are given in terms of the function $f(z)$ by

$$\sum_{v=1}^n a_v = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(z) \frac{e^{\omega z}}{z} dz, \quad (8)$$

where $c > \max(S, 0)$, $\lambda_n < \omega < \lambda_{n+1}$, and the integration contour does not pass through $\{\lambda_n\}$. If $\omega = \lambda_n$, then the term a_n in the sum of the left-hand side of (8) is replaced by $a_n/2$ (O. Perron, 1908). Furthermore, if $S < x$, then we have

$$a_n = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{x_0}^{x_0+T} f(x + iy) \exp[\lambda_n(x + iy)] dy \quad (9)$$

(J. Hadamard, 1908; C. Tanaka, 1952).

If $x = \operatorname{Re} z > S$, then $f(z) = o(|y|)$ ($|y| \rightarrow \infty$). In order to investigate its behavior more precisely, Bohr introduced

$$\mu(x) = \limsup_{|y| \rightarrow +\infty} \frac{\log |f(x + iy)|}{\log |y|}$$

in his thesis (1910) and called it the **order** over $\operatorname{Re} z = x$. The function $\mu(x)$ is nonnegative, monotone decreasing, convex, and continuous with respect to x . Bohr later found that there is

a kind of periodicity for the values of $f(z)$ over $\operatorname{Re} z = x$; this was the origin of the theory of \dagger almost periodic functions.

As for the zeros of the function $f(z)$, the following theorems are known: If $f(z)$ is not identically zero, it has only a finite number of zeros in $x \geq S + \varepsilon$, $e^{-Mx} \leq y \leq e^{Mx}$ for arbitrary positive numbers ε, M (Perron, 1908). If we denote by $N(T)$ the number of zeros in $x > S + \varepsilon$, $T < y < T + 2\delta \log T$, then $\limsup_{T \rightarrow \infty} N(T)/(\log T)^2 \leq \delta/\varepsilon$ (E. Landau, 1927).

There have been many investigations into the connection between the singularities of $f(z)$ and the coefficients a_n . If the a_n are real and positive, the point $z = S$ is always a singularity of $f(z)$. Moreover, if $S = 0$, $\operatorname{Re} a_n \geq 0$, and $\lim_{n \rightarrow \infty} (\cos(\arg a_n))^{1/\lambda_n} = 1$, then $z = 0$ is a singularity of $f(z)$ (C. Biggeri, 1939). Furthermore, if $\lambda_n/n \rightarrow \infty$, $\liminf_{n \rightarrow \infty} (\lambda_{n+1} - \lambda_n) > 0$, then the line $\operatorname{Re} z = S$ is the \dagger natural boundary of $f(z)$ (F. Carleson and Landau, 1921; A. Ostrowski, 1923). If $S = 0$, $\liminf_{n \rightarrow \infty} (\lambda_{n+1} - \lambda_n) = q > 0$, then there always exist singularities on every interval on the imaginary axis with the length $2\pi/q$ (G. Pólya, 1923). S. Mandelbrojt (1954, 1963) gave some interesting results concerning the relations between the singularities of (1) and the Fourier transform of an entire function.

If $U = -\infty$, the function $f(z)$ is an entire function. Its \dagger order (in the sense of entire function) ρ is given by

$$\rho = \limsup_{x \rightarrow -\infty} \frac{\log^+ \log^+ M(x)}{|x|},$$

$$M(x) = \sup_{-\infty < y < \infty} |f(x + iy)|.$$

There have been many investigations into the \dagger Julia direction of $f(z)$ and related topics by Mandelbrojt, Valiron, and Tanaka.

D. Tauberian Theorems

As in the case of power series, if the series $\sum a_n$ converges to s , then $f(+0) = s$ (**Abel's continuity theorem**). The converse is not necessarily true. The converse theorems, with additional conditions on a_n and λ_n , are called **Tauberian theorems**, as in the case of power series. Many theorems are known about this field. The most famous additional conditions are $\lim_{n \rightarrow \infty} \lambda_n a_n / (\lambda_n - \lambda_{n-1}) = 0$ (Landau, 1926) and $a_n = O((\lambda_n - \lambda_{n-1})/\lambda_n)$ (K. Ananda-Rau, 1928).

For the summation of Dirichlet series (especially \dagger Riesz's method of summation) \rightarrow 379 Series R. For Tauberian theorems (especially the Wiener-Ikehara-Landau theorem) of the ordinary Dirichlet series \rightarrow 123 Distribution of Prime Numbers B.

E. Series Related to Dirichlet Series

A series of the form

$$\sum_{n=1}^{\infty} \frac{n! a_n}{z(z+1)(z+2)\dots(z+n)}, \quad z \neq 0, -1, -2, \dots$$

is called a **factorial series** with the coefficients $\{a_n\}$. It converges or diverges simultaneously with the ordinary Dirichlet series $\sum a_n/n^z$ except at $z=0$ and negative integers. The series

$$\sum_{n=1}^{\infty} \frac{(z-1)(z-2)\dots(z-n)}{n!} a_n = \sum_{n=1}^{\infty} a_n \binom{z-1}{n}$$

is called a **binomial coefficient series**. It converges or diverges simultaneously with the ordinary Dirichlet series $\sum (-1)^n a_n/n^z$ except at $z=0$ and positive integers.

References

- [1] V. Bernshteĭn, Leçons sur les progrès récents de la théorie des séries de Dirichlet, Gauthier-Villars, 1933.
- [2] K. Chandrasekharan and S. Minakshisundaram, Typical means, Oxford Univ. Press, 1952.
- [3] G. H. Hardy and M. Riesz, The general theory of Dirichlet's series, Cambridge, 1915.
- [4] S. Mandelbrojt, Séries lacunaires, Actualités Sci. Ind., Hermann, 1936.
- [5] S. Mandelbrojt, Séries adhérentes, régularisation des suites, applications, Gauthier-Villars, 1952.
- [6] S. Mandelbrojt, Dirichlet series, Reidel, 1972.
- [7] G. Valiron, Théorie générale des séries de Dirichlet, Mém. Sci. Math., Gauthier-Villars, 1926.

122 (IV.14)
Discontinuous Groups

A. Definitions [1-4]

Suppose that a group Γ acts continuously on a $^*\text{Hausdorff}$ space X , that is, for every $\gamma \in \Gamma$ and $x \in X$, an element γx of X is assigned in such a way that the mapping $x \rightarrow \gamma x$ is a homeomorphism of X onto itself and that we have $\gamma_1(\gamma_2 x) = (\gamma_1 \gamma_2)x$, $1x = x$, where 1 is the identity element of Γ . Two points $x, y \in X$ are said to be Γ -**equivalent** if there exists a $\gamma \in \Gamma$ such that $y = \gamma x$. (Γ -equivalence for subsets of X is defined similarly.)

We consider the following conditions of discontinuity of Γ . (i) For every $x \in X$ and any

infinite sequence $\{\gamma_i\}$ consisting of distinct elements of Γ , the sequence $\{\gamma_i x\}$ has no $^*\text{cluster point}$ in X . (ii) For every $x \in X$, there exists a neighborhood U_x such that $\gamma U_x \cap U_x = \emptyset$ for all but finitely many $\gamma \in \Gamma$. (ii') If $x, y \in X$ are not Γ -equivalent, there exist neighborhoods U_x, U_y of x, y , respectively, such that $\gamma U_x \cap U_y = \emptyset$ for all $\gamma \in \Gamma$. (iii) For any compact subset M of X , $\gamma M \cap M = \emptyset$ for all but finitely many $\gamma \in \Gamma$.

It is easy to see that (ii) \Rightarrow (i), (ii) + (ii') \Rightarrow (iii); if, moreover, X is $^*\text{locally compact}$, we also have (iii) \Rightarrow (ii), (ii'). When (i) holds, Γ is called a **discontinuous transformation group** of X , and when (ii) holds, Γ is called a **properly discontinuous transformation group**. In particular, when X can be identified with a $^*\text{homogeneous space}$ G/K of a locally compact group G by a compact subgroup K , the conditions (i), (ii), and (iii) for a subgroup Γ of G are all equivalent, and they are also equivalent to the condition that Γ is a $^*\text{discrete subgroup}$ of G .

For a discontinuous group Γ acting on X , the $^*\text{stabilizer}$ $\Gamma_x = \{\gamma \in \Gamma \mid \gamma x = x\}$ of $x \in X$ is always a finite subgroup. When $\Gamma_x = \{1\}$ for all $x \in X$, Γ is said to be **free** (or to **act freely** on X). If $\Gamma_x = \bigcap_{x \in X} \Gamma_x = \{1\}$, Γ is said to act $^*\text{effectively}$ on X . A point $x \in X$ is called a **fixed point** of Γ if $\Gamma_x \neq \Gamma_x$. In the following, we assume for simplicity that Γ acts effectively on X , unless otherwise specified.

Since Γ -equivalence is clearly an $^*\text{equivalence relation}$, we can decompose X into Γ -equivalence classes, or Γ - $^*\text{orbits}$. The space of all Γ -orbits, called the **quotient space** of X by Γ , is denoted by $\Gamma \backslash X$. When Γ satisfies the conditions (ii) and (ii'), the space $\Gamma \backslash X$ becomes a $^*\text{Hausdorff space}$ with respect to the topology of the quotient space. If, moreover, Γ is free, X is an (unramified) $^*\text{covering space}$ of $\Gamma \backslash X$ with the $^*\text{covering transformation group}$ Γ . (Conversely, a covering transformation group is always a free, properly discontinuous transformation group.) In general, X may be viewed as a covering space of $\Gamma \backslash X$ with ramifications, and the ramifying points (in X) are nothing but the fixed points of Γ .

B. Fundamental Regions

A complete set of representatives F of $\Gamma \backslash X$ in X (that is, a subset F of X such that $\Gamma F = X$, $\gamma F \cap F = \emptyset$ for $\gamma \in 1, \gamma \neq 1$) is called a **fundamental region** of Γ in X if it further satisfies suitable topological or geometrical requirements. Here we assume that \bar{F} , the closure of F , is the closure of its interior F^i . (In this case, \bar{F} or F^i is sometimes called the **fundamental region** of Γ instead of F itself.) Such a fundamental region exists if Γ satisfies the

conditions (ii), (ii'), and the set of fixed points is \dagger nowhere dense in X (R. Baer, F. W. Levi, 1931). A fundamental region F is called **normal** if the set $\{\gamma F\} (\gamma \in \Gamma)$ is locally finite, that is, if, for every $x \in X$, there exists a neighborhood U_x such that $\gamma F \cap U_x = \emptyset$ for all but a finite number of $\gamma \in \Gamma$. If X is \dagger connected and F is normal, then Γ is generated by the set of $\gamma \in \Gamma$ such that $\gamma \bar{F} \cap \bar{F} \neq \emptyset$. Thus it is useful to have a fundamental region in order to find a set of generators of Γ and a set of \dagger fundamental relations for them. When X has a Γ -invariant \dagger Borel measure μ and Γ is countable, then the measure $\mu(F)$ of F is independent of the choice of F . Hence it is legitimate to put $\mu(\Gamma \backslash X) = \mu(F)$; Γ is called a **discontinuous group of the first kind** (C. L. Siegel [2]) if Γ is a discontinuous transformation group which has a normal fundamental region F such that $\{\gamma | \gamma \bar{F} \cap \bar{F} = \emptyset\}$ is finite and $\mu(F) < \infty$. For instance, if X is locally compact and \bar{F} is compact ($\Leftrightarrow \Gamma \backslash X$: compact), then Γ is of the first kind.

When we are concerned only with the qualitative properties of Γ , it is sometimes convenient to relax the conditions for a fundamental region, replacing it by a **fundamental (open) set** Ω of Γ , that is, an (open) subset Ω of X such that $\Gamma \Omega = X$ and $\gamma \Omega \cap \Omega = \emptyset$ for all but a finite number of $\gamma \in \Gamma$ [5-9].

C. The Case of a Riemann Surface

Let Γ be a discontinuous group of analytic automorphisms of a \dagger Riemann surface X . In virtue of \dagger uniformization theory, it is enough, in principle, to consider the case where X is \dagger simply connected. Thus we have the following three cases:

(1) $X = \mathbb{C} \cup \{\infty\}$ (\dagger Riemann sphere). Γ is a finite group. Since Γ can also be considered as a group of motions of the sphere, it is either a cyclic, \dagger dihedral, or \dagger regular polyhedral group [10].

(2) $X = \mathbb{C}$ (complex plane). Γ is contained in the group of motions of the plane. The subgroup consisting of all parallel translations contained in Γ is a \dagger free Abelian group of rank $v \leq 2$. If $v = 0$, then Γ is a finite cyclic group. When $v > 0$, Γ consists of the transformations of the following form:

When $v = 1$, $z \rightarrow e^{k\omega} z + m\omega$ $(k, m \in \mathbb{Z})$,

When $v = 2$, $z \rightarrow e^{k\omega} z + m_1 \omega_1 + m_2 \omega_2$ $(k, m_i \in \mathbb{Z})$,

where $\omega, \omega_1, \omega_2$ are nonzero complex numbers with $\text{Im}(\omega_2/\omega_1) > 0$, and $\varepsilon = \pm 1$ in general, except in special cases when $v = 2$ and $\omega_2/\omega_1 = \zeta_4$ (resp. ζ_3 or ζ_6 , where $\zeta_l = \exp(2\pi i/l)$), in

which cases we may put $\varepsilon = \zeta_4$ (resp. ζ_3 or ζ_6). For the fundamental regions corresponding to these values of ε , see Fig. 1. In the cases $v = 1$ and 2, the \dagger automorphic functions with respect to Γ are essentially given by exponential functions and elliptic functions, respectively (\rightarrow 134 Elliptic Functions).

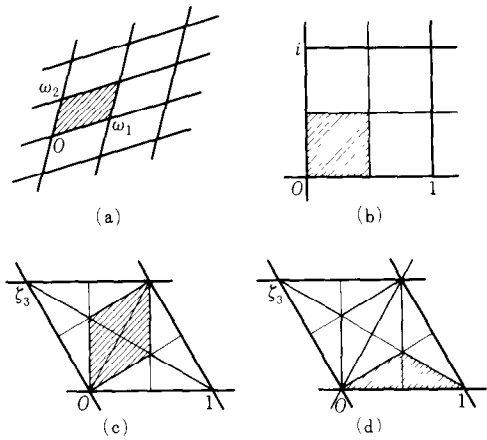


Fig. 1
(a) $v = 2, \varepsilon = 1$. (b) $\omega_1 = 1, \omega_2 = i, \varepsilon = i$. (c) $\omega_1 = 1, \omega_2 = \zeta_3, \varepsilon = \zeta_3$. (d) $\omega_1 = 1, \omega_2 = \zeta_3, \varepsilon = \zeta_6$.

(3) $X = \{|z| < 1\}$ (unit disk) [3, 10, 11]. By a \dagger Cayley transformation, the unit disk can be transformed to the upper half-plane $\mathfrak{H} = \{z = x + iy | y > 0\}$. Any analytic automorphism of \mathfrak{H} is given by a real \dagger linear fractional transformation (Möbius transformation) $z \rightarrow (az + b)/(cz + d)^{-1}$ ($a, b, c, d \in \mathbb{R}, ad - bc = 1$). The totality of real linear fractional transformations acts transitively on \mathfrak{H} . Hence \mathfrak{H} can be identified with the \dagger homogeneous space G/K of $G = SL(2, \mathbb{R})$ by $K = SO(2)$ (which is the stabilizer of the point $\sqrt{-1}$). Hence discontinuous groups Γ of analytic automorphisms of \mathfrak{H} are obtained as discrete subgroups of G . Actually, every element of G defines an analytic automorphism of the whole Riemann sphere, which leaves the real axis $\mathbb{R} \cup \{\infty\}$ invariant. For any $z \in \mathbb{C} \cup \{\infty\}$ and a sequence $\{\gamma_i\}$ consisting of distinct elements of Γ , a cluster point of the sequence $\{\gamma_i z\}$ in $\mathbb{C} \cup \{\infty\}$ is called a **limit point** of Γ . When only one or two limit points exist, Γ can easily be transformed to one of the groups given in (2). Otherwise, the set \mathfrak{Q} of all limit points of Γ is infinite, and either $\mathfrak{Q} = \mathbb{R} \cup \{\infty\}$ or \mathfrak{Q} is a \dagger perfect, \dagger nowhere dense subset of $\mathbb{R} \cup \{\infty\}$. When \mathfrak{Q} is infinite, Γ is called a **Fuchsoid group**.

Since \mathfrak{H} has a G -invariant \dagger Riemannian metric $ds^2 = y^{-2}(dx^2 + dy^2)$ (called the \dagger Poincaré metric), by which \mathfrak{H} becomes a **hyperbolic plane** (\dagger non-Euclidean plane with negative curvature), we can construct a fundamental region F of Γ which is a normal polygon bounded by geodesics, that is, the arcs of cir-

cles orthogonal to the real axis. A set of generators of Γ and the fundamental relations for them are easily obtained by observing the correspondence of the equivalent sides of the fundamental polygon. (Conversely, starting from a normal polygon satisfying a suitable condition, one can construct a discontinuous group Γ having F as a fundamental region. In this manner, we generally obtain a (nontrivial) continuous family of discrete subgroups of G .)

A Fuchsoid group Γ is finitely generated if and only if the fundamental polygon F has a finite number of sides, and in that case Γ is called a **Fuchsian group**. (More generally, a finitely generated discontinuous group of linear fractional transformations acting on a domain in the complex plane is called a **Kleinian group**.) A Fuchsian group Γ is of the **first kind** if and only if $\mathfrak{Q} = \mathbf{R} \cup \{\infty\}$; otherwise, it is of the **second kind**. It is also known that a discontinuous group Γ is a Fuchsian group of the first kind if and only if $\mu(\Gamma \backslash \mathfrak{H}) < \infty$ [12]. For a real point $x \in \mathbf{R} \cup \{\infty\}$, we also denote by Γ_x the stabilizer of x (in Γ). The point x is called a (**parabolic**) **cusp** of Γ if Γ_x is a free cyclic group generated by a \dagger parabolic transformation ($\neq \pm 1$). Cusps of Γ are represented by vertices of the fundamental polygon on the real axis. On the other hand, if a fixed point z of Γ lies in \mathfrak{H} , then the stabilizer Γ_z is always a finite cyclic group generated by an \dagger elliptic transformation. Hence such a point z is also called an **elliptic point** of Γ . For a Fuchsian group Γ of the first kind, let $\{z_1, \dots, z_s\}$ be a complete set of representatives of the Γ -equivalence classes of the elliptic points of Γ (which can also be chosen from among the vertices of the fundamental polygon), and let e_i be the order of Γ_{z_i} ; furthermore, let t be the number of the Γ -equivalence classes of parabolic cusps of Γ . Then the quotient space $\Gamma \backslash \mathfrak{H}$ can be compactified by adjoining t points at infinity, and the resulting space becomes a compact Riemann surface \mathfrak{R}_Γ if we define an analytic structure on it in a suitable manner. The area $\mu(\mathfrak{R}_\Gamma)$ measured by the Poincaré metric is given by the \dagger Gauss-Bonnet formula:

$$\mu(\mathfrak{R}_\Gamma) = \int_F \frac{dx dy}{y^2} \\ = 2\pi \left[2g - 2 + \sum_{i=1}^s \left(1 - \frac{1}{e_i} \right) + t \right],$$

where g is the \dagger genus of the Riemann surface \mathfrak{R}_Γ . It is known that there exists a lower bound ($= \pi/21$) for $\mu(\mathfrak{R}_\Gamma)$ [12]. Automorphic functions (or Fuchsian functions) with respect to a Fuchsian group Γ , which are essentially the same thing as algebraic functions on the Riemann surface \mathfrak{R}_Γ , have been objects of extensive study since Poincaré (1882).

D. Modular Groups [10, 13]

The group

$$\Gamma = SL(2, \mathbf{Z}) \\ = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| a, b, c, d \in \mathbf{Z}, ad - bc = 1 \right\}$$

(or the corresponding group of linear fractional transformations) is called the (**elliptic**) **modular group**. The modular group Γ is a Fuchsian group of the first kind acting on \mathfrak{H} , and its fundamental region together with the correspondence of the equivalent sides is shown in Fig. 2. Fig. 3 illustrates the transformations under Γ of the fundamental triangle, where Γ is regarded as acting on the unit disk. From Fig. 2 we obtain the generators of $\Gamma \pmod{\{\pm I_2\}}$ ($I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$) and the

fundamental relations:

$$\sigma_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2^2 = (\sigma_2 \sigma_1)^3 = -I_2.$$

There are two Γ -equivalence classes of elliptic points of Γ , which are represented by $\zeta_4 = i$ and ζ_3 , with $[\Gamma_i : \{\pm I_2\}] = 2$, $[\Gamma_{\zeta_3} : \{\pm I_2\}] = 3$; and only one Γ -equivalence class of parabolic cusps, which coincides with $\mathbf{Q} \cup \{\infty\}$. The corresponding Riemann surface \mathfrak{R}_Γ is analytically equivalent to the Riemann sphere.

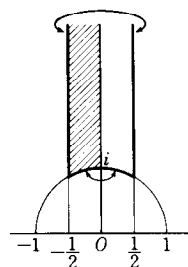


Fig. 2

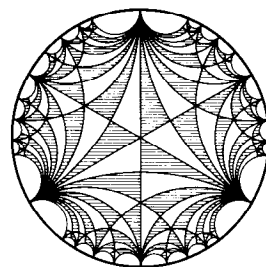


Fig. 3

For a positive integer N , the totality $\Gamma(N)$ of elements in Γ satisfying the condition $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N}$ forms a normal subgroup of Γ , called a **principal congruence subgroup of level N** . (For the case $N = 2$, see Fig. 4.) In general, a subgroup Γ' of Γ containing $\Gamma(N)$ for some N is called a **congruence subgroup** of Γ . (It is known that there actually exists a subgroup Γ' of Γ with a finite index, which is not a congruence subgroup.) For $N \geq 3$, $-I_2 \notin \Gamma(N)$, so that $\Gamma(N)$ is effective. (For $N = 1, 2$, we have $\Gamma(N)_5 = \{\pm I_2\}$.) If $N \geq 2$, $\Gamma(N)$ has no elliptic point. The number $t(N)$ of the equivalence classes of cusps of $\Gamma(N)$ and the genus $g(N)$ of the corresponding Rie-

mann surface $\mathfrak{H}_{\Gamma(N)}$ are given as follows:

$$t(1)=1, \quad t(2)=3,$$

$$t(N)=(1/2N)[\Gamma:\Gamma(N)] \quad (N\geq 3),$$

$$g(1)=g(2)=0,$$

$$g(N)=1+((N-6)/24N)[\Gamma:\Gamma(N)] \quad (N\geq 3),$$

where $[\Gamma:\Gamma(N)]=N^3\prod_{p|N}(1-1/p^2)$. Automorphic functions with respect to $\Gamma(N)$ are called \dagger modular functions of level N .

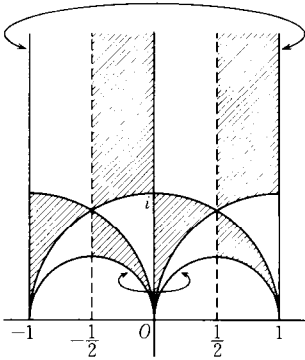


Fig. 4
A fundamental region of $\Gamma(2)$ that consists of six fundamental regions of $\Gamma(1)$.

E. The Case of Many Variables

Up to the present time, discontinuous groups Γ and the corresponding automorphic functions have been studied only in the following cases: (2') $X=\mathbb{C}^n$, $\Gamma\cong\mathbb{Z}^{2n}$ (the free Abelian group of rank $2n$, consisting of parallel translations) [14] (\rightarrow 3 Abelian Varieties); (3') X is a bounded domain in \mathbb{C}^n and Γ is a discontinuous group of analytic automorphisms of X . (In this case, conditions (i), (ii), (iii) are equivalent.)

In the case (3'), the group $\mathcal{A}(X)$ of all (complex) analytic automorphisms of X , endowed with its natural (\dagger compact-open) topology, becomes a \dagger Lie group, which has Γ as a discrete subgroup. When $\Gamma\backslash X$ is compact, it is known by the theory of automorphic functions (or by a theorem of Kodaira) that $\Gamma\backslash X$ becomes a \dagger projective variety, which is a \dagger minimal model [3, 14]. In particular, when X is a \dagger symmetric bounded domain, i.e., when X becomes a \dagger symmetric Riemannian space with respect to its \dagger Bergman metric, the connected component G of the identity element of $G'=\mathcal{A}(X)$ (which incidentally coincides with that of the group $I(X)$ of all \dagger isometries of X) is a \dagger semisimple Lie group of noncompact type (i.e., without compact simple factors), and X can be identified with the homogeneous space of G by a maximal compact subgroup K of G . The theory of discontinuous groups of this type, initiated by Siegel (especially in the case

where $X=\mathfrak{H}_n=Sp(n,\mathbb{R})/K$, \dagger Siegel's upper half-space; $\Gamma=Sp(n,\mathbb{Z})$, \dagger Siegel's modular group of degree n), O. Blumenthal, H. Braun, and L.-K. Hua, and continued by those in the German school such as M. Koecher, H. Maass, and others, has undergone substantial development in recent years under the influence of the theory of algebraic groups [5, 8, 14–17] (\rightarrow 32 Automorphic Functions).

On the other hand, for a symmetric Riemannian space X of negative curvature, the group of isometries $G=I(X)$ is a \dagger semisimple Lie group of noncompact type with a finite number of connected components and with a finite center, and X can be identified with the homogeneous space of G by a maximal compact subgroup. Therefore the study of discontinuous groups of isometries of X can be reduced to that of discrete subgroups of a Lie group G of this type. A typical example is the case where X is the space of all real \dagger positive definite symmetric matrices of degree n with determinant 1; this space can be identified with the quotient space $SL(n,\mathbb{R})/SO(n)$ ($A\in SL(n,\mathbb{R})$ acts on X by $X\ni S\rightarrow^{-1}ASA$). The unimodular group $\Gamma=SL(n,\mathbb{Z})$ is a discontinuous group of the first kind acting on this space X , and a method of constructing a fundamental region of Γ in X is provided by the **Minkowski reduction theory** [6, 7].

F. Discrete Subgroups of a Semisimple Lie Group

Two subgroups Γ, Γ' of a group G are called **commensurable** if $\Gamma\cap\Gamma'$ is of finite index in both Γ and Γ' . For a real \dagger linear algebraic group $G\subset GL(n,\mathbb{R})$ defined over \mathbb{Q} , a subgroup Γ commensurable with $G_{\mathbb{Z}}=G\cap GL(n,\mathbb{Z})$ is called an **arithmetic subgroup** (in the original sense) of G (examples: $SL(n,\mathbb{Z}), Sp(n,\mathbb{Z})$). An arithmetic subgroup Γ is always discrete, and when G is semisimple, the quotient space $\Gamma\backslash G$ is of finite volume ($\mu(\Gamma\backslash G)<\infty$) with respect to an invariant measure μ . Moreover, $\Gamma\backslash G$ is compact if and only if G is $\dagger\mathbb{Q}$ -compact (or $\dagger\mathbb{Q}$ -anisotropic), that is, if $G_{\mathbb{Q}}$ or $G_{\mathbb{Z}}$ consists of only \dagger semisimple elements (A. Borel, Harish-Chandra, G. D. Mostow, and T. Tamagawa [6, 7]); the same results remain true if G is \dagger Zariski connected and has no \dagger character defined over \mathbb{Q} . The proofs of these facts (and the compactification of the quotient space $\Gamma\backslash X$ for the noncompact case) depend on a construction of fundamental open sets that generalizes the reduction theory of Minkowski and Siegel [5, 8, 15–18].

For a connected semisimple Lie group G of noncompact type and a discrete subgroup Γ with $\mu(\Gamma\backslash G)<\infty$, the following **density**

theorem holds (Borel, *Ann Math.*, (2) 72 (1960)):

(i) For any linear representation ρ of G , the linear closure of $\rho(\Gamma)$ coincides with that of $\rho(G)$; (ii) if G is algebraic, Γ is t Zariski dense in G . Furthermore, suppose that G is a direct product of simple groups G_i and the center of G is finite; Γ is called **irreducible** if its projection on any (proper) partial product of $\{G_i\}$ is not discrete. For instance, if G is a $^t\mathbf{Q}$ -simple algebraic group then $\Gamma = G_{\mathbf{Z}}$ is irreducible. In general, there exists a partition of the set of indices $\{i\}$ such that Γ is commensurable with a direct product of irreducible discrete subgroups of the partial products corresponding to this partition, and these irreducible components are unique up to commensurability.

The method of constructing a discrete subgroup Γ of $G = SL(2, \mathbf{R})$ in a geometric manner using the upper half-plane can be generalized to some extent to the construction of discrete subgroups of certain groups using hyperbolic spaces of low dimensions (E. B. Vinberg). Except for these few cases, today it is known that a discrete subgroup Γ of a semisimple Lie group G (of \mathbf{R} -rank ≥ 2) with $\mu(\Gamma \backslash G) < \infty$ is arithmetic in a certain sense (\rightarrow Section G). This implies there are only very few discrete subgroups for a semisimple Lie group of higher rank. Actually a number of facts suggesting this result were already known in the 1960s. First, the only subgroups of $SL(n, \mathbf{Z})$ ($n \geq 3$), $Sp(n, \mathbf{Z})$ ($n \geq 2$) with finite index are congruence subgroups (H. Bass, M. Lazard, and J.-P. Serre; this result has been generalized to the case of an arbitrary t Chevalley group over an algebraic number field by C. C. Moore and H. Matsumoto). Second, $G_{\mathbf{Z}} = SL(n, \mathbf{Z})$, $Sp(n, \mathbf{Z})$ are maximal in G [8]. Finally, it is known (rigidity theorem) that if a connected semisimple Lie group G with a finite center does not contain a simple factor which is locally isomorphic to $SL(2, \mathbf{R})$, then any discrete subgroup Γ of G with compact quotient $\Gamma \backslash G$ has no nontrivial t deformation (i.e., all deformations are obtained from inner automorphisms of G) (A. Selberg, E. Calabi and E. Vesentini, and A. Weil [19]). This last result amounts to the vanishing of the cohomology group $H^1(\Gamma, X, \text{ad})$, and in this connection an extensive study has been made by Y. Matsushima, S. Murakami, G. Shimura, and K. G. Ramanan to determine the t Betti numbers of $\Gamma \backslash X$, and more generally the cohomology groups of the type $H(\Gamma, X, \rho)$ with an arbitrary representation ρ of G . These cohomology groups are closely related to automorphic forms with respect to Γ [8, 20].) For a t nilpotent or t solvable Lie group, a general method of constructing discrete subgroups is known; see, for example, M. Saito, *Amer. J. Math.*, 83 (1961).

G. Rigidity and Arithmeticity

A discrete subgroup Γ of a Lie group G with $\mu(\Gamma \backslash G) < \infty$ is usually called a **lattice** of G . A lattice Γ of G is said to be **uniform** if $\Gamma \backslash G$ is compact. By a theorem of Borel and Harish-Chandra [7], an arithmetic subgroup of a real linear group (defined over \mathbf{Q}) is a lattice. Using this result, Borel showed further, by a constructive method, that any semisimple Lie group has a lattice, especially a uniform one (Borel, *Topology*, 2 (1963)).

For a long time there were no known examples of nonarithmetic irreducible lattices in semisimple Lie groups other than those locally isomorphic to $SL_2(\mathbf{R})$. This naturally led to Selberg's conjecture that any irreducible (nonuniform) lattice in a semisimple Lie group G not locally isomorphic to $SL_2(\mathbf{R})$ is arithmetic (Selberg, International Colloquium on Function Theory, Bombay, 1960). The conjecture seemed to be well-grounded by the rigidity theorem of Weil and Selberg [19].

However, in 1966–1967, V. S. Makarov and Vinberg constructed nonarithmetic non-uniform lattices in $SO(n, 1)$ ($n = 3, 4, 5$) by a geometric method; the lattices are generated by t reflections [21]. Thus rigidity and arithmeticity do not necessarily coincide, and the conjecture should be considered under stronger conditions.

As for rigidity of uniform lattices, Mostow established in 1970 the following **strong rigidity theorem** [22]: If G, G' are semisimple Lie groups with trivial center and without compact factors, and are not locally isomorphic to $SL_2(\mathbf{R})$, and if Γ, Γ' are irreducible uniform lattices, then any isomorphism $\theta: \Gamma \rightarrow \Gamma'$ extends to an analytic isomorphism $\tilde{\theta}: G \rightarrow G'$ (namely, $\tilde{\theta}|_{\Gamma} = \theta$). The previous rigidity theorem is now implied by Mostow's. If X is a simply connected symmetric space, then a t locally symmetric space M covered by X is expressed as a quotient of X by a fixed-point-free properly discontinuous group Γ in the group $I(X)$ of total isometries (when the Lie algebra of $I(X)$ does not have a compact factor, this condition for Γ is equivalent to saying that it is a torsion-free discrete subgroup of the semisimple Lie group $I(X)$), and the fundamental group $\pi_1(M)$ of $M = \Gamma \backslash X$ is isomorphic to Γ . The strong rigidity theorem implies that compact locally symmetric spaces M_1 and M_2 (of higher dimensions) are isomorphic as Riemannian spaces if and only if $\pi_1(M_1)$ and $\pi_1(M_2)$ are isomorphic as abstract groups.

On the other hand, in 1973, G. A. Margulis proved the arithmeticity of irreducible non-uniform lattices in semisimple real linear groups of \mathbf{R} -rank greater than 1 (*Russian Math. Surveys*, 29 (1974) (original in Russian,

1974); *Functional Anal. Appl.* 9 (1975) (original in Russian, 1975). M. S. Raghunathan also proved independently the same fact under a slightly stronger condition. Their results, together with the rigidity of nonuniform lattices in (higher-dimensional) semisimple Lie groups of \mathbf{R} -rank 1 established by H. Garland, Raghunathan, and G. Prasad (*Inventiones Math.*, 21 (1973)), imply that the strong rigidity theorem holds similarly for nonuniform lattices.

The results of Margulis and Raghunathan show that the Selberg conjecture in the original sense is affirmative for the case of \mathbf{R} -rank greater than 1. But neither argument is applicable to uniform lattices, for they depend deeply on the fact (proved by D. A. Kazdan and Margulis) that a nonuniform lattice contains a nonidentity unipotent element. Previously, in a lecture at the international congress of mathematicians at Moscow, 1966, I. I. Pyatetskii-Shapiro generalized the definition of arithmeticity and suggested that arithmeticity of lattices should be investigated without the distinction of whether they are uniform or nonuniform. His definition is equivalent to the following [9, 24]: For a connected semisimple algebraic group G defined over \mathbf{R} , a subgroup $\Gamma \subset G = G_{\mathbf{R}}$ is an **arithmetic subgroup** (of G) if there is an algebraic group H defined over \mathbf{Q} and a surjective homomorphism $\varphi: H \rightarrow \text{Ad } G$ defined over \mathbf{R} such that the Lie group $(\text{Ker } \varphi)_{\mathbf{R}}$ is compact and that $\varphi(H_{\mathbf{Z}})$ and $\text{Ad } \Gamma$ are commensurable. The uniform lattice in G that is constructed by the method of Borel is arithmetic in this sense. In 1974, Margulis finally established the following arithmeticity theorem [23, 24]: If the \mathbf{R} -rank of G is not less than 2, an irreducible lattice Γ in G is an arithmetic subgroup of G (even if it is uniform). In the same lecture, Pyatetskii-Shapiro also extended the Selberg conjecture to such "semisimple Lie groups" as those containing p -adic Lie groups as factors. Margulis proved this Pyatetskii-Shapiro conjecture affirmatively by showing that an analog of the strong rigidity theorem holds for such groups.

As for semisimple Lie groups of \mathbf{R} -rank 1, besides the lattices constructed by Makarov and Vinberg there are only the few examples of nonarithmetic lattices in $SU(2, 1)$ presented by Mostow (*Proc. Nat. Acad. Sci. US*, 75 (1978); *Pacific J. Math.*, 86 (1980)). The problem of arithmeticity still remains open for groups locally isomorphic to $SO(n, 1)$ ($n \geq 6$), $SU(n, 1)$ ($n \geq 3$), $Sp(n, 1)$, or F_4 .

H. Geometric Discontinuous Groups [1, 25]

The study of discontinuous groups Γ acting on a Euclidean or projective space X as a transformation group of a given structure is a class-

Discontinuous Groups

ical problem. All possibilities for such Γ have been enumerated in low-dimensional cases. For instance, there are 230 kinds of discontinuous groups of \dagger Euclidean motions acting on 3-dimensional Euclidean space without fixed subspaces, which are classified into 32 \dagger crystal classes (A. Schönflies and E. S. Fedorov, 1891–1892; \rightarrow 92 Crystallographic Groups). All discontinuous groups of a Euclidean space generated by \dagger reflections have also been enumerated (H. S. M. Coxeter, 1934 [25, 26]).

I. Kleinian Groups

The last decade has seen considerable research on (finitely generated) Kleinian groups. This research is closely related to the theory of \dagger quasiconformal mappings and the \dagger moduli of Riemann surfaces.

Making use of Eichler cohomology and \dagger potentials, L. V. Ahlfors established his finiteness theorem and L. Bers his area theorem. Bers and B. Maskit investigated the boundaries of \dagger Teichmüller spaces and discovered Kleinian groups with the property that the complement of the set \mathfrak{L} of limit points is connected and simply connected.

Numerous mathematicians have subsequently discussed the classification, deformation, and stability properties of the set \mathfrak{L} , uniformization and deformation of Riemann surfaces with or without nodes, and other geometric properties. In their discussions, the theory of quasiconformal mappings has played an important role. The discontinuous groups of motions of hyperbolic 3-space have also been studied.

References

- [1] B. L. van der Waerden, Gruppen von linearen Transformationen, *Erg. Math.*, Springer, 1935 (Chelsea, 1948).
- [2] C. L. Siegel, Discontinuous groups, *Ann. Math.*, (2) 44 (1943), 674–689 (*Gesammelte Abhandlungen*, Springer 1966, vol. 2, 390–405).
- [3] Sémin. H. Cartan 6, Fonctions automorphes et espaces analytiques, *Ecole Norm. Sup.*, 1953–1954.
- [4] G. Shimura, Introduction to the arithmetic theory of automorphic functions, *Publ. Math. Soc. Japan*, 11 (1971).
- [5] Sémin. H. Cartan 10, Fonctions automorphes, *Ecole Norm. Sup.*, 1957–1958.
- [6] A. Weil, Discontinuous subgroups of classical groups, *Lecture notes*, Univ. of Chicago, 1958.
- [7] A. Borel and Harish-Chandra, Arithmetic subgroups of algebraic groups, *Ann. Math.*, (2) 75 (1962), 485–535.

- [8] Algebraic groups and discontinuous subgroups, Amer. Math. Soc. Proc. Symp. Pure Math., 1966, vol. 9.
- [9] M. S. Raghunathan, Discrete subgroups of Lie groups, Erg. Math., Springer, 1972.
- [10] R. Fricke and F. Klein, Vorlesungen über die Theorie der automorphen Funktionen I, Teubner, 1897, second edition, 1926.
- [11] H. Poincaré, Théorie des groupes fuchs-
iens, Acta Math., 1 (1882), 1–62 (Oeuvres,
Gauthier-Villars, 1916, vol. 2, 108–168).
- [12] C. L. Siegel, Some remarks on discontinu-
ous groups, Ann. Math., (2) 46 (1945), 708–718
(Gesammelte Abhandlungen, Springer, 1966,
vol. 3, 67–77).
- [13] F. Klein and R. Fricke, Vorlesungen über
die Theorie der elliptischen Modulfunktionen
I, Teubner, 1890.
- [14] C. L. Siegel, Analytic functions of several
complex variables, Lecture notes, Institute for
Advanced Study, Princeton, 1948–1949.
- [15] C. L. Siegel, Symplectic geometry, Amer.
J. Math., 65 (1943), 1–86 (Gesammelte Ab-
handlungen, Springer, 1966, vol. 2, 274–359;
Academic Press, 1964).
- [16] I. Satake, On compactifications of the
quotient spaces for arithmetically defined
discontinuous groups, Ann. Math., (2) 72
(1960), 555–580.
- [17] W. L. Baily and A. Borel, Compactifi-
cations of arithmetic quotients of bounded
symmetric domains, Ann. Math., (2) 84 (1966),
442–528.
- [18] H. Minkowski, Diskontinuitätsbereich
für arithmetische Äquivalenz, J. Reine Angew.
Math., 129 (1905), 220–274 (Gesammelte Ab-
handlungen, Teubner, 1911, vol. 2, 53–100;
Chelsea, 1967).
- [19] A. Weil, On discrete subgroups of Lie
groups, Ann. Math., (2) 72 (1960), 369–384; II,
ibid., 75 (1962), 578–602.
- [20] Y. Matsushima and S. Murakami, On
vector bundle valued harmonic forms and
automorphic forms on symmetric Riemannian
manifolds, Ann. Math., (2) 78 (1963), 365–416.
- [21] E. B. Vinberg, Discrete groups generated
by reflections in Lobacevskii spaces, Math.
USSR-Sb., 1 (1967), 429–444. (Original in
Russian, 1967.)
- [22] G. D. Mostow, Strong rigidity of locally
symmetric spaces, Ann. Math. Studies 78,
Princeton Univ. Press, 1973.
- [23] G. A. Margulis, Discrete groups of mo-
tions of manifolds of nonpositive curvature,
Amer. Math. Soc. Transl., (2) 109 (1977), 33–
45. (Original in Russian, 1974.)
- [24] G. A. Margulis, Arithmetic irreducible
lattices in semisimple groups of rank greater
than 1 (in Russian), Mir, 1977. (Appendix to
the Russian translation of [9].)
- [25] H. S. M. Coxeter and W. O. J. Moser,

Generators and relations for discrete groups,
Erg. Math., Springer, 1957.

[26] N. Bourbaki, Eléments de mathématique,
Groupes et algèbres de Lie, ch. 4, 5, 6, Ac-
tualités Sci. Ind., 1377, Hermann, 1968.

[27] L. Bers and I. Kra, A crash course on
Kleinian groups, Lecture notes in math. 400,
Springer, 1974.

123 (V.7) Distribution of Prime Numbers

A. General Remarks

Given a real number x , we denote by $\pi(x)$ the
number of primes not exceeding x . A. M.

Legendre (1808) obtained empirically the for-
mula $\pi(x) \approx x/(\log x - B)$ for some constant B ,
and C. F. Gauss (1849) obtained the formula

$$\pi(x) \approx \int_2^x \frac{du}{\log u},$$

assuming the average density of primes to be
 $1/\log x$. The **Bertrand conjecture**, which asserts
the existence of at least one prime between x
and $2x$, was proved by P. L. Chebyshev (1848),
who introduced the functions

$$\theta(x) = \sum_{m=1}^{\infty} \sum_{p \leq x} \log p$$

and

$$\begin{aligned} \psi(x) &= \sum_{p^m \leq x} \log p \\ &= \theta(x) + \theta(\sqrt{x}) + \theta(\sqrt[3]{x}) + \dots \end{aligned}$$

(In this section, p represents a prime number.)
He thereby proved

$$Ax + O(\sqrt{x}) < \theta(x) < \psi(x) < (6/5)Ax + O(\sqrt{x}),$$

where $A = \log(2^{1/2} 3^{1/3} 5^{1/5} 30^{-1/30})$. G. F. B.
Riemann (1858) considered the function $\zeta(s)$
(where $s = \sigma + it$ is a complex variable), ex-
pressed by the Dirichlet series $\sum_{n=1}^{\infty} n^{-s}$, which
is convergent for $\sigma > 1$. He found relations
between the zeros of $\zeta(s)$ (\rightarrow 450 Zeta Func-
tions) and $\pi(x)$. F. Mertens (1874) obtained the
formulas

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + O(1),$$

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + B + O\left(\frac{1}{\log x}\right),$$

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) = \frac{e^{-c}}{\log x} \left(1 + O\left(\frac{1}{\log x}\right)\right),$$

where c is the Euler constant and B is some
constant.

B. Prime Number Theorem

The prime number theorem

$$\lim_{x \rightarrow \infty} \frac{\pi(x) \log x}{x} = 1, \quad \text{or} \quad \pi(x) \sim \frac{x}{\log x},$$

was proved almost simultaneously (1896) by J. Hadamard and C. J. de La Vallée-Poussin. Without using the theory of entire functions, E. Landau (1908) established the formula

$$\pi(x) = \text{Li } x + O(xe^{-c\sqrt{\log x}}),$$

where

$$\text{Li } x = \lim_{\delta \rightarrow +0} \left(\int_0^{1-\delta} + \int_{1+\delta}^x \right) \frac{du}{\log u}$$

is the †logarithmic integral. It can be shown by integration by parts that

$$\begin{aligned} \text{Li } x &= \frac{x}{\log x} + \frac{1!x}{\log^2 x} + \dots \\ &\quad + \frac{(k-1)!x}{\log^k x} + O\left(\frac{x}{\log^{k+1} x}\right). \end{aligned}$$

For example, by taking $x = 10^7$, we get $\pi(x) = 664,579$, $\text{Li } x \approx 664,918$, and $x/\log x \approx 620,417$.

If the Dirichlet series $f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ satisfies the condition $\sum_{n \leq x} a_n \sim cx$, then its abscissa of convergence is 1, and we have $\lim_{s \rightarrow 1+0} (s-1)f(s) = c$. The converse is known as the †Tauberian theorem. If $F(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ ($a_n \geq 0$) converges absolutely for $\sigma > 1$ and $F(s) - c/(s-1)$ is analytic for $\sigma \geq 1$, then we obtain $\sum_{n \leq x} a_n \sim cx$ (**Wiener-Ikehara-Landau theorem**, 1932). Specifically, if we put $-\zeta'(s)/\zeta(s) = \sum_{n=1}^{\infty} \Lambda(n)n^{-s}$, then the conditions of the theorem are satisfied, and we obtain $\sum_{n \leq x} \Lambda(n) = \psi(x) \sim x$.

It is easily seen that the prime number theorem is equivalent to $\psi(x) \sim x$ or $\theta(x) \sim x$. The †number-theoretic function $\Lambda(n)$ (**Mangoldt's function**) introduced above satisfies $\sum_{d|n} \Lambda(d) = \log n$. It follows from the †Möbius inversion formula that $\Lambda(n) = \sum_{d|n} \mu(d) \log(n/d)$. Hence $\Lambda(n) = \log p$ ($n = p^l$) and $= 0$ otherwise. Thus we obtain $\psi(x) = \sum_{n \leq x} \Lambda(n) = O(x)$. When $f(x) = \theta(x)$ or $\psi(x)$, it is easy to show that $\int_2^x (f(t)/t^2) dt = \log x + O(1)$ and $\liminf_{x \rightarrow \infty} f(x)/x \leq 1 \leq \limsup_{x \rightarrow \infty} f(x)/x$. However, it is not easy to prove $f(x) \sim x$. To do so, introduce the number-theoretic function $M(n)$, which satisfies $\sum_{d|n} M(d) = \log^2 n$. As before, we have $M(n) = \sum_{d|n} \mu(d) \log^2(n/d)$; hence

$$M(n) = \begin{cases} (2l-1) \log p & (n = p^l, l \geq 1), \\ 2 \log p \log q & (n = p^l q^m, l \geq 1, m \geq 1), \\ 0 & (\text{otherwise}). \end{cases}$$

Thus we obtain $\sum_{n \leq x} M(n) = 2x \log x + O(x)$. This leads to A. Selberg's well-known formula

(1949):

$$\theta(x) \log x + \sum_{p \leq x} \theta(x/p) \log p = 2x \log x + O(x),$$

which enabled him to prove $\theta(x) \sim x$. Thus he obtained for the first time a proof of the prime number theorem that does not use complex analytic methods. The simple formulas $\sum_{n=1}^{\infty} \mu(n)/n = 0$ and $\sum_{n \leq x} \mu(n) = o(x)$, obtained by H. von Mangoldt (1897), were revealed by Landau to have a deep meaning concerning the prime number theorem. Let $\pi_r(x)$ denote the number of integers not exceeding x that can be expressed as the product of r distinct primes. In generalizing the prime number theorem, Landau (1911) proved that

$$\pi_r(x) \sim \frac{1}{(r-1)!} \frac{x(\log \log x)^{r-1}}{\log x}.$$

Let us write $\mathfrak{G}(x) = \sum_{n=1}^{\infty} e^{-\pi x n^2}$. Riemann proved that

$$\begin{aligned} &\Gamma\left(\frac{s}{2}\right) \pi^{-s/2} \zeta(s) \\ &= \frac{1}{s(s-1)} + \int_1^{\infty} \mathfrak{G}(x) (x^{s/2-1} + x^{-1/2-s/2}) dx \end{aligned}$$

and obtained the well-known functional equation for the zeta function (\rightarrow 450 Zeta Functions B)

$$\pi^{-s/2} \Gamma(s/2) \zeta(s) = \pi^{-1/2+s/2} \Gamma(1/2-s/2) \zeta(1-s).$$

This enables us to extend $\zeta(s)$ as a meromorphic function to the whole complex plane. Utilizing this extended $\zeta(s)$ and the following result of O. Perron on Dirichlet series, we can estimate $\psi(x)$. Let $\sigma_0 (\neq \infty)$ be the abscissa of convergence of $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$, and let $a > 0$, $a > \sigma_0$, and $x > 0$. If

$$\lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{a-iT}^{a+iT} F(s) \frac{x^s}{s} ds$$

exists, then the limit is equal to $\sum'_{n \leq x} f(n)$, where \sum' means that in the summation the last term $f(x)$ is replaced by $f(x)/2$ if x is an integer. In many cases, $F(s)$ has a pole at $s = 1$, and the principal part of the sum is obtained from the residue at $s = 1$, whereas the residual part is given by a certain contour integral. To estimate $\psi(x)$, we use $-\zeta'(s)/\zeta(s)$ as $F(s)$; hence the problem arises of determining the zeros of $\zeta(s)$. Riemann conjectured that all the zeros of $\zeta(s)$ in the strip $0 \leq \sigma \leq 1$ must be situated on the vertical line $\sigma = 1/2$. If this so-called †Riemann hypothesis (\rightarrow 450 Zeta Functions) is true, then it follows that $\pi(x) = \text{Li } x + O(\sqrt{x} \log x)$. The ultimate validity of Riemann's hypothesis remains in doubt. Concerning this, the most recent major result is the following formula, obtained by I. M. Vinogradov (1958): $\pi(x) = \text{Li } x +$

$O(x \exp(-c \log^{3/5} x / \log \log^{1/5} x))$. Without using Riemann's hypothesis, J. E. Littlewood (1918) proved that

$$\limsup_{x \rightarrow \infty} \frac{\pi(x) - \text{Li } x}{(\sqrt{x} / \log x) \log \log \log x} > 0,$$

$$\liminf_{x \rightarrow \infty} \frac{\pi(x) - \text{Li } x}{(\sqrt{x} / \log x) \log \log \log x} < 0.$$

If we denote by $N(T)$ the number of zeros of $\zeta(s)$ in the domain $0 < \sigma < 1$, $0 < t < T$, then we have

$$N(T) = \frac{1}{2\pi} T \log T - \frac{1 + \log 2\pi}{2\pi} T + O(\log T).$$

Let $N_0(T)$ denote the number of zeros of $\zeta(s)$ on the interval $\sigma = 1/2$, $0 < t < T$. Selberg (1942) obtained the impressive result

$$N_0(T) > cT \log T.$$

E. C. Titchmarsh (1936) showed that there exist 1041 zeros of $\zeta(s)$ in the domain $0 < \sigma < 1$, $0 < t < 1468$ and that all lie on the line $\sigma = 1/2$. Computers have provided further results that seem to justify the Riemann hypothesis. For example, it has been calculated that there are 75,000,000 zeros of $\zeta(s)$ in the domain $0 < \sigma < 1$, $0 < t < 32,585,736.4$ and that all are simple zeros and lie on the line $\sigma = 1/2$ (R. P. Brent, *Math. Comp.*, 33 (1979)). N. Levinson proved in 1974 by another method that at least one-third of the zeros of the Riemann zeta function are on the line $\sigma = 1/2$. The minimum of the modulus of the imaginary part of the zeros with $\sigma = 1/2$ is $t = 14.13 \dots$

C. Twin Primes

Let p_n be the n th prime. We know from the prime number theorem that $p_n \sim n \log n$; more precisely, $p_n = n \log n + n \log \log n + O(n)$. A pair of primes differing only by 2 are called **twin primes**. It is still unknown whether there exist infinitely many twin primes. There exist infinitely many n satisfying $p_{n+1} - p_n < c \log p_n$ (P. Erdős, 1940). Suppose that $\zeta(1/2 + it) = O(|t|^c)$. A. E. Ingham (1937) proved that

$$p_{n+1} - p_n < p_n^\Theta, \quad \Theta = (1 + 4c)/(2 + 4c) + \varepsilon,$$

by using the following density theorem related to the zeros of $\zeta(s)$: If we denote by $N(\alpha, T)$ the number of zeros of $\zeta(s)$ in the domain $\alpha \leq \sigma \leq 1$, $0 \leq t \leq T$ ($1/2 \leq \alpha \leq 1$), then there is a positive constant c such that $N(\alpha, T) = O(T^{2(1+2c)(1-\alpha)} \log^5 T)$. The **Lindelöf hypothesis** asserts that the constant c can be made arbitrarily small. If the Riemann hypothesis holds, then the Lindelöf hypothesis also holds. It is clear that we can substitute $1/6 + \varepsilon$ ($\varepsilon > 0$) for c , ε being arbitrarily small. W. Haneke

(1963) showed that c can be replaced by $6/37 + \varepsilon$, and consequently that $\Theta \leq 61/98 + \varepsilon$. D. R. Heath-Brown and H. Iwaniec (1979) proved that $\Theta \leq 11/20 + \varepsilon$ by using the \dagger sieve method and the zero density theorem of L -functions (Section E). R. A. Rankin (1935) proved that

$$p_{n+1} - p_n > c \log p_n \log \log p_n \log \log \log p_n \times (\log \log \log p_n)^{-2}$$

holds for infinitely many n . If we denote by $\pi_2(x)$ the number of primes $p \leq x$ such that $p + 2$ is also a prime, then it has been conjectured by Hardy and Littlewood (*Acta Math.*, 44 (1922)) that

$$\pi_2(x) \sim C \int_2^x \frac{du}{\log^2 u}$$

as $x \rightarrow \infty$, where

$$C = 2 \prod_{p>2} \left\{ 1 - \frac{1}{(p-1)^2} \right\} = 1.32032 \dots$$

The numerical evidence provided by the computation $\pi_2(10^9) = 3424506$ (Brent, *Math. Comp.*, 28 (1974)) tends to indicate the truth of this conjecture. At present, $76 \cdot 3^{139} \pm 1$ seems to be the largest known pair of twin prime numbers (H. C. Williams and C. R. Zarnke, *Math. Comp.*, 26 (1972)).

D. Prime Numbers in Arithmetic Progressions

Let k be a positive integer, $\chi(n)$ be a residue character modulo k (\rightarrow 295 Number-Theoretic Functions D), and $L(s, \chi) = \sum_{n=1}^{\infty} \chi(n)n^{-s}$ ($\sigma > 1$) be the \dagger Dirichlet L -function. The function $L(s, \chi)$ of s defined by this series can be extended to an analytic function in the whole complex plane in the same way as the Riemann zeta function. In particular, when χ is the principal character, then $L(s, \chi)$, thus extended, is a meromorphic function whose only pole is situated at $s = 1$ and is simple; otherwise the function $L(s, \chi)$ is holomorphic on \mathbb{C} .

Using this function $L(s, \chi)$ and in connection with his research concerning the \dagger class numbers of quadratic forms, P. G. L. Dirichlet (1837) proved that there exist infinitely many primes in the arithmetical progression $l, l+k, l+2k, \dots$, where l is the initial term and k a common difference relatively prime to l . This result is called the **Dirichlet theorem** (or **prime number theorem for arithmetic progressions**).

Suppose that α runs over all integral ideals in a \dagger quadratic number field K of discriminant d . Then the \dagger Dedekind zeta function $\zeta_K(s)$ of K is defined by $\sum (N\alpha)^{-s}$ for $\sigma > 1$. By virtue of the decomposition law of prime ideals (\rightarrow 347 Quadratic Fields C), we have $\zeta_K(s) = \zeta(s)L(s, \chi)$, where $\chi(n) = (d/n)$ is the \dagger Kronecker symbol.

Utilizing $\zeta_K(s)$, we obtain formulas concerning the \dagger class number $h(d)$ of the field K . If $d > 0$, then $h(d) = (\sqrt{d/2\log \varepsilon})L(1, \chi)$, where ε is the \dagger fundamental unit of K . On the other hand, if $d < 0$, then $h(d) = (w\sqrt{-d/2\pi})L(1, \chi)$, where w denotes the number of the roots of unity contained in K . It follows that $L(1, \chi) \geq 2\log((1 + \sqrt{5}/2)/\sqrt{|d|})$. Let χ be a character modulo k induced by a primitive character χ^0 . Since we have $L(s, \chi) = L(s, \chi^0) \prod_{p|k} (1 - \chi^0(p)p^{-s})$, it can be shown that $L(1, \chi) \neq 0$ for a real character χ . It is easy to prove that $L(1, \chi) \neq 0$ for a complex character χ . These statements then lead to the Dirichlet theorem. The proof was simplified by H. N. Shapiro (1951). Besides Landau's three proofs for $L(1, \chi) \neq 0$ for real character χ (1908), there are elegant proofs by T. Estermann (1952), Selberg (1949), and others. For a character $\chi \bmod k$, we always have $L(1, \chi) = O(\log k)$, while $L(1, \chi)^{-1} = O(\log k)$ with one possible exception, which may occur only if χ is a real character. Even in this case, we have $L(1, \chi)^{-1} = O(k^\varepsilon)$ (where $\varepsilon > 0$ is arbitrary, but O depends on ε). This result was obtained by K. L. Siegel (1934) from his study concerning class numbers of imaginary quadratic number fields. His proof was simplified by Estermann (1948) and S. D. Chowla (1950). The importance of the prime number theorem for arithmetic progressions was revealed when it was applied to the Goldbach problem (\rightarrow 4 Additive Number Theory C). Concerning this problem, the manner in which the remainder term depends on the modulus k became an object of investigation. The Page-Siegel-Walfisz theorem is convenient to use: Denote by $\pi(x; k, l)$ the number of primes not exceeding x and of the form $ky + l$, where $(k, l) = 1$. If $x \geq \exp(k^\varepsilon)$ (where $\varepsilon > 0$ is arbitrary), then we have

$$\pi(x; k, l) = \frac{\text{Li } x}{\varphi(k)} + O\left(\frac{x e^{-c(\varepsilon)\sqrt{\log x}}}{\varphi(k)}\right).$$

Further research on the distribution of zeros of $L(s, \chi)$ is necessary for the study of $\pi(x; k, l)$ when x takes smaller values. If χ is a nonprincipal real character, then $L(s, \chi)$ may have at most one real zero β_1 around 1; this is called **Siegel's zero**. Because of this fact, when x is small we are unable to obtain any formula to indicate the uniform distribution of primes. However, we have the following deep result, obtained by E. Fogels (1962). For a given positive ε , there exist $c_0(\varepsilon)$ and $c(\varepsilon)$ such that $\pi(x; k, l) > c(\varepsilon)x/\varphi(k)k^\varepsilon \log x$, provided that $x \geq k^{c_0(\varepsilon)}$. On the other hand, Titchmarsh (1930), using the \dagger sieve method, obtained $\pi(x; k, l) = O(x/\varphi(k) \log x)$ for $x \geq k^{c_0}$. A theorem of this type is called the **Burn-Titchmarsh theorem** (\rightarrow Section E). Fogels's theorem is based on the following theorem by Yu. V. Linnik (1947) and

K. Prachar (1957), which is an extension of Page's theorem (1935): We let δ be the function defined by $\delta = 1 - \beta_1$ if $L(s, \chi)$ has an exceptional real zero β_1 , and $\delta = c_1/\log(k(|t| + 2))$ otherwise (where c_1 is a suitably small number). If we denote by β the real part of any zero ($\neq \beta_1$) of $L(s, \chi)$, then the theorem states that

$$\beta \leq 1 - \frac{c_1}{\log(k(|t| + 2))} \log\left(\frac{c_1 e}{\delta \log(k(|t| + 2))}\right),$$

provided that $\delta \log(k(|t| + 2)) \leq c_1$. Linnik called this result the **Deuring-Heilbronn phenomenon**. Using this theorem and the zero density theorem, Linnik (1944) proved that $p(k, l) \ll k^L$, where $p(k, l)$ is the least prime in the arithmetic progression $l, l + k, l + 2k, \dots$ and L is a constant. We call L **Linnik's constant**. M. Jutila (1977) and Chen Jing-Run (1979) proved that $L \leq 80$ and $L \leq 17$, respectively.

Let s be positive, b_j, z_j ($1 \leq j \leq s$) be complex, and l, m be real numbers satisfying $\max |z_j| \geq 1$, $l \geq s$, and $m \geq 0$. Under these conditions P. Turán (1953) obtained the following fundamental theorem, which is called the **power sum theorem**:

$$\max_{m \leq r \leq l+m} |b_1 z_1^r + \dots + b_s z_s^r| \geq \left(\frac{l}{8e(l+m)}\right)^l \min_{1 \leq j \leq s} |b_1 + \dots + b_j|.$$

This theorem is effective in research on the distribution of zeros of zeta functions. Based on this new method, Turán (1961), S. Knapowski (1962), and Fogels (1965) reached the results cited above.

Another method of research, considered to be a new sieve method, on the distribution of primes was introduced by Selberg and A. I. Vinogradov. This method was followed by W. B. Jurkat and H. E. Richert (1965). Linnik, A. Rényi (1950), and E. Bombieri (1965) founded still another method, called the **large sieve method**, by which P. T. Batemann, Chowla, and P. Erdős studied the value of $L(1, \chi)$.

E. Sieve Method

Let A be a set of integers, and P a set of primes. For each $p \in P$, let Ω_p be a set of residues mod p , and $\omega(p)$ the number of residues belonging to Ω_p . The sieve method is a device for estimating (from above or below) the number of integers n belonging to the set $S(A, P, \Omega_p) = \{n | n \in A, n \bmod p \notin \Omega_p \text{ for } p \in P\}$. The combinatorial methods of the Brun, Busstab, and Richert sieves are interesting and efficient but quite complicated. Here, we shall briefly describe the **Selberg sieve**. As an example, denote by $S(x; q, l)$ the number of n satisfying $n \equiv l \bmod q$, $n \leq x$, $(n, D) = 1$, where q

is a prime not exceeding x , $z \leq x$, $D = \prod_{p \leq z} p$, and $(l, q) = 1$; then

$$S(x; q, l) = \sum_{\substack{n \leq x \\ n \equiv l \pmod{q}}} \sum_{d|(n, D)} \mu(d) \leq \sum_{\substack{n \leq x \\ n \equiv l \pmod{q}}} \left(\sum_{d|(n, D)} \lambda_d \right)^2,$$

where $\lambda_1 = 1$ and λ_d is arbitrary for $d > 1$. Thus the problem reduces to the optimization of λ_d . Proceeding in this manner, C. Hooley (1967) and Y. Motohashi (1975), using analytic methods, obtained certain deep results relating to the Brun-Titchmarsh theorem. Using the large sieve, H. L. Montgomery and R. C. Vaughan (1973) expressed this theorem in a precise form: $\pi(x, q, l) < 2x/\varphi(q) \log(x/q)$ for all $q < x$.

Let n_1, n_2, \dots, n_Z be Z natural numbers not exceeding N , and $Z(p, a)$ the number of n_j 's such that $n_j \equiv a \pmod{p}$. Rényi (1950) proved that

$$\sum_{p \leq \sqrt[3]{N/12}} p \sum_{a=1}^{p-1} \left\{ Z(p, a) - \frac{Z}{p} \right\}^2 \leq 2NZ.$$

Set $a_n = 1$ for $n = n_j$ and $a_n = 0$ otherwise, and set $S(\alpha) = \sum_{n \leq N} a_n \exp(2\pi i n \alpha)$; then

$$p \sum_{a=1}^{p-1} \left\{ Z(p, a) - \frac{Z}{p} \right\}^2 = \sum_{a=1}^{p-1} \left| S\left(\frac{a}{p}\right) \right|^2.$$

In view of this simple fact, Bombieri (1965) and P. X. Gallagher (1968) extended the problem and proved that, in general,

$$\sum_{q \leq Q} \sum_{\substack{a=1 \\ (a, q)=1}}^q \left| \sum_{\substack{M < n \leq M+N \\ a_n \exp\left(2\pi i \frac{a}{q} n\right)}} a_n \right|^2 \leq (N + 2Q^2) \sum_{M < n \leq M+N} |a_n|^2.$$

Similar results can be obtained for character sums $\sum_{M < n \leq M+N} a_n \chi(n)$. In this connection Montgomery proved that

$$S(\{n; M < n \leq M+N\}; P, \Omega_p) \leq (N + 2Q^2) \left\{ \sum_{q \leq Q} \prod_{p|q} \frac{\omega(p)}{p - \omega(p)} \right\}^{-1},$$

where

$$P(z) = \prod_{\substack{p \in P \\ p \leq z}} p.$$

Using these methods, the following estimate was obtained by Bombieri:

$$\sum_{q \leq x^{1/2}(\log x)^{-B}} \max_{y \leq x} \max_{\substack{l \\ (q, l)=1}} \left| \pi(y, q, l) - \frac{1}{\varphi(q)} \int_2^y \frac{du}{\log u} \right| \ll x(\log x)^{-A},$$

where A is arbitrary and B is a certain func-

tion of A (Vaughan). These methods, known collectively as the **large sieve**, were first directed toward proving the **Rényi theorem** stating that every sufficiently large even integer can be represented as the sum of a prime and an almost prime integer (M. B. Barban). Afterward, combining Richert's sieve with this large sieve, Chen Jing-Run (1973) improved this result to a remarkable degree (\rightarrow 4 Additive Number Theory C). Several applications of Bombieri's theorem have been demonstrated by P. D. T. A. Elliot and H. Halberstam (1966): e.g., the estimation of the number of representations of n as $p + x^2 + y^2$ (Hooley, Linnik) and the estimation of $\sum_{p \leq n} d(n-p)$ (Linnik, B. M. Bredihin).

Let $N(\alpha, T, \chi)$ denote the number of zeros of $L(s, \chi)$ in the rectangle $\alpha \leq \sigma \leq 1$, $|t| \leq T$. Combining the large sieve with new Fourier integral techniques and the Turán † power-sum method, Gallagher (1970) proved that there exists a positive constant c satisfying

$$\sum_{\chi \pmod{Q}} N(\alpha, T, \chi) \ll (QT)^{c(1-\alpha)},$$

$$\sum_{q \leq Q} \sum_{\chi \pmod{q}} N(\alpha, T, \chi) \ll (QT)^{c(1-\alpha)}.$$

Similar results were also obtained by G. Halasz and Montgomery (1969) using another method, which was further exploited by Jutila, M. H. Huxley, and Iwaniec. In particular, Heath-Brown and Iwaniec (1979) deduced that $\pi(x+y) - \pi(x) \geq cy(\log x)^{-1}$ if $y \geq x^{1/20+\varepsilon}$. Combined with the Deuring-Heilbronn phenomenon, the zero density theorem above not only establishes the Linnik theory, but also yields the following result, due to K. A. Rodoskiĭ, T. Tatzuzaawa, and A. I. Vinogradov:

$$\pi(x; q, l) = \frac{1}{\varphi(q)} \int_2^x \frac{du}{\log u} - E \frac{\chi_1(l)}{\varphi(q)} \int_2^x \frac{u^{\beta-1}}{\log u} du + O\left(\frac{1}{\varphi(q)} x^{1-\frac{c}{\Delta}}\right)$$

if $x \geq \exp(\log q \log \log q)$, where $E = 1$ or 0 according as Siegel's zero β exists or not, and where

$$\Delta = \text{Max}(\log q, (\log x)^{2/5} (\log \log x)^{1/5}).$$

Throughout these researches, estimates of the type

$$\sum_{\chi \pmod{q}} \left| L\left(\frac{1}{2} + it, \chi\right) \right|^4 \ll \varphi(q) \{|t| + 2\} \log^c \{q(|t| + 2)\}$$

and

$$\sum_{\chi \pmod{q}} \int_{-T}^T \left| L\left(\frac{1}{2} + it, \chi\right) \right|^4 dt \ll \varphi(q) (T+2) \log^c q(T+2)$$

are of great importance and have been studied by A. F. Lavrik (1968), Linnik (1961), Huxley (1972), and K. Ramachandra (1975).

F. The Prime Ideal Theorem in Algebraic Number Fields

In an algebraic number field of finite degree, the prime number theorem is replaced by the **prime ideal theorem** (T. Mitsui, 1956, Fogels, 1962), which is based on the theory of the Hecke $\dagger L$ -function (E. Hecke, 1917; Landau, 1918). Let K be a finite Galois extension over an algebraic number field k of finite degree. Suppose that \mathfrak{p} is a prime ideal of k and is not ramified in K . The \dagger Frobenius automorphism of a prime divisor of \mathfrak{p} in K determines a conjugate class C of the Galois group of K/k . Let $\pi(x; C)$ denote the number of prime ideals in k associated with the class C in the above sense and whose norm does not exceed x . Then we have

$$\pi(x, C) = \frac{h(C)}{[K:k]} \text{Li } x + O(xe^{-c\sqrt{\log x}}),$$

where $h(C)$ is the number of elements contained in C and c is a positive constant depending on K/k . This is an extension of Chebotarev's theorem (E. Artin, 1923).

References

- [1] R. Ayoub, An introduction to the analytic theory of numbers, Amer. Math. Soc. Math. Surveys, 1963.
- [2] H. Bohr and H. Cramér, Die neuere Entwicklung der analytischen Zahlentheorie, Enzykl. Math., II C (8), 1922.
- [3] K. Chandrasekharan, Introduction to analytic number theory, Springer, 1968.
- [4] P. Erdős, Some recent advances and current problems in number theory, Lectures on Modern Mathematics III, T. L. Saaty (ed.), Wiley, 1965, 196–244.
- [5] E. Hecke, Vorlesungen über die Theorie der algebraischen Zahlen, Akademische Verlag, Leipzig, 1923 (Chelsea, 1970).
- [6] E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen I, II, Teubner, 1909 (Chelsea, 1969).
- [7] E. Landau, Vorlesungen über Zahlentheorie I, II, Hirzel, 1927 (Chelsea, 1946).
- [8] G. H. Hardy, Divergent series, Clarendon Press, 1949.
- [9] H. M. Edwards, Riemann's zeta function, Academic Press, 1974.
- [10] A. E. Ingham, The distribution of prime numbers, Cambridge Univ. Press, 1932.

- [11] D. N. Lehmer, List of prime numbers from 1 to 10,006,721, Carnegie Institution of Washington, 1914.
- [12] N. Levinson, More than one third of zeros of Riemann's zeta-function on $\sigma = 1/2$, Advances in Math., 13 (1974), 383–436.
- [13] K. Prachar, Primzahlverteilung, Springer, 1957.
- [14] A. Selberg, An elementary proof of the prime number theorem for arithmetic progressions, Canad. J. Math., 2 (1950), 66–78.
- [15] E. C. Titchmarsh, The theory of the Riemann zeta-function, Clarendon Press, 1951.
- [16] I. M. Vinogradov, A new estimate of the function $\zeta(1+it)$ (in Russian), Izv. Akad. Nauk SSSR, 22 (1958), 161–164.
- [17] M. N. Huxley, The distribution of prime numbers, Oxford Univ. Press, 1972.
- [18] P. Turán, Eine neue Methode in der Analysis und deren Anwendungen, Akadémiai Kiadó, Budapest, 1953.
- [19] H. L. Montgomery, Topics in multiplicative number theory, Lecture notes in math., 227, Springer, 1971.
- [20] E. Bombieri, Le grand crible dans la théorie analytique des nombres, Société Mathématique de France, 1974.
- [21] H. Halberstam and H. E. Richert, Sieve methods, Academic Press, 1974.
- [22] C. Hooley, Applications of sieve methods to the theory of numbers, Cambridge Univ. Press, 1976.
- [23] H. E. Richert, Lectures on sieve methods, Tata Inst., 1976.

124 (XI.8) Distribution of Values of Functions of a Complex Variable

A. General Remarks

Suppose that we are given a function $f: A \rightarrow B$ and that the variables z, w take on values in A, B , respectively. A **value distribution** of $f(z)$ is a set of points z where $f(z)$ takes on a certain value w (called w -points of $f(z)$). Value distribution theory is usually concerned with the study of value distributions of complex \dagger analytic functions. Value distribution theory has been developed extensively and deeply for the case where A is the finite plane $|z| < \infty$ or the unit disk $|z| < 1$ and B is the extended complex plane $|w| \leq \infty$, and there are many interesting results in this case (\rightarrow 272 Meromorphic Functions). Value distributions for analytic functions on general domains or on Riemann

surfaces or of several complex variables have also been studied.

B. The Case of $|z| < R \leq \infty$

For a †transcendental entire function $f(z)$, every value (including ∞) is a value of the †cluster set of $f(z)$ at the point at infinity (†Weierstrass's theorem). This theorem was improved in the following way by E. Picard in 1879: A transcendental entire function $f(z)$ has an infinite number of w -points for any finite value w except for at most one finite value (†Picard's theorem). A value w for which the w -points are at most finite is said to be a Picard's exceptional value. E. Borel gave a precise form of this theorem, taking into consideration the order of a function, and G. Julia proved the existence of Julia's directions (\rightarrow 272 Meromorphic Functions; 429 Transcendental Entire Functions). After other results in value distribution theory had been obtained by J. Hadamard, G. Valiron, and others, R. Nevanlinna published an important work in 1925 in which he established the so-called **Nevanlinna theory** of meromorphic functions in $|z| < R \leq \infty$, unifying results obtained up until that time, and which became the starting point of the subsequent value distribution theory (\rightarrow 272 Meromorphic Functions). T. Shimizu and L. V. Ahlfors gave a geometric meaning to the Nevanlinna †characteristic function $T(r)$. Ahlfors established the theory of covering surfaces by metricotopological methods in 1935, and applied it to obtain the Nevanlinna theory and many other results on meromorphic functions. This theory revealed that the topological meaning of the number 2 of Picard's exceptional values is closely related to the †Euler characteristic 2 of the sphere. H. Selberg established the value distribution theory of †algebroidal functions and gave a precise form of G. Rémoundos's theorem, which corresponds to Picard's theorem for algebroidal functions.

Moreover, as an extension of the value distribution theory of meromorphic functions, there is the theory of holomorphic curves or of systems of entire functions. Let f_0, f_1, \dots, f_n ($n \geq 1$) be entire functions without common zeros for all and for which $f_0 : f_1 : \dots : f_n$ is not constant. Put $f = (f_0, f_1, \dots, f_n)$. This is a reduced representation of a nonconstant holomorphic curve $f: C \rightarrow P^n(C)$. For $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in C^{n+1} - \{0\}$, considering the zeros of

$$(\alpha, f) = \alpha_0 f_0 + \alpha_1 f_1 + \dots + \alpha_n f_n \quad (\neq 0),$$

we can extend Picard's theorem and the Nevanlinna theory for meromorphic functions to holomorphic curves.

We define the characteristic function $T(r) (= T(r, f))$ as

$$T(r) = \frac{1}{2\pi} \int_0^{2\pi} \log |f| d\theta \Big|_{r_0}^r,$$

$$|f| = (|f_0|^2 + |f_1|^2 + \dots + |f_n|^2)^{1/2},$$

where r_0 is a fixed positive number and $A(r)|_{r_0}^{r_0} = A(r) - A(r_0)$. Using a bivector $[ff'] = (f_i f'_j - f_j f'_i)$ of f and $f' = (f'_0, f'_1, \dots, f'_n)$, we have another representation of $T(r)$:

$$T(r) = \frac{1}{\pi} \int_{r_0}^r \frac{ds}{s} \int_0^s \int_0^{2\pi} \frac{|[ff']|^2}{|f|^2} t dt d\theta.$$

f is transcendental when $\lim T(r)/\log r = \infty$, and the number

$$\lambda = \dim \{(c_0, c_1, \dots, c_n) \in C^{n+1} \mid c_0 f_0 + c_1 f_1 + \dots + c_n f_n = 0\}$$

is the degeneracy index of f . It holds that $0 \leq \lambda \leq n-1$. As an extension of Picard's theorem, J. Dufresnoy stated that, for transcendental holomorphic curves f and among α in general position, the zeros of (α, f) are infinite except for at most $n + \lambda + 1$. X ($\subset C^{n+1} - \{0\}$) is in general position when any p ($\leq n+1$) vectors in X are linearly independent. In connection with this result, there are many Picard-type theorems, and when $\lambda > 0$, there are some particular results for holomorphic curves. H. Cartan, H. and J. Weyl, and Ahlfors extended the Nevanlinna theory to holomorphic curves as follows. For $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in C^{n+1} - \{0\}$, we put

$|\alpha|$ = the length of α ,

$$|\alpha f| = |(\alpha, f)|/|\alpha||f|, \quad (\alpha, f) \neq 0,$$

and define the proximity function

$$m(r, \alpha) = \frac{1}{2\pi} \int_0^{2\pi} \log \frac{1}{|\alpha f|} d\theta$$

and the counting function

$$N(r, \alpha) = \frac{1}{2\pi} \int_0^{2\pi} \log |(\alpha, f)| d\theta \Big|_{r_0}^r.$$

Then we have the first main theorem: For any α for which $(\alpha, f) \neq 0$,

$$T(r) + m(r_0, \alpha) = N(r, \alpha) + m(r, \alpha);$$

and the second main theorem: When $\lambda = 0$, for any $\alpha_1, \alpha_2, \dots, \alpha_q$ in general position,

$$(q - n - 1)T(r) < \sum_{j=1}^q N(r, \alpha_j) + S(r),$$

where $S(r) = O(\log r T(r))$ except for r in a set e of finite logarithmic measure, (i.e., $\int_e d \log r < \infty$). For $\lambda > 0$, it is conjectured that “ $q - n - 1$ ” can be changed into “ $q - n - \lambda - 1$.” This is unsolved except for some special cases.

Some results show relations between exceptional values and the order as in the case of meromorphic functions. Nevanlinna theory has been extended to the associated curves f^p ($p = 1, 2, \dots, n; f^1 = f$) for $\lambda = 0$, and there have been attempts to extend this theory of holomorphic curves even further by generalizing domains or ranges.

C. The Case of General Domains

The value distributions of meromorphic functions defined in a general domain or an open Riemann surface depend on the function-theoretic "size" of the set of singularities (\rightarrow 169 Function-Theoretic Null Sets) or the type of the Riemann surface (\rightarrow 367 Riemann Surfaces). For instance, we have the following theorem of Picard type: A single-valued meromorphic function with a set of singularities of \dagger logarithmic capacity zero takes on every value infinitely often in any neighborhood of each singularity except for at most an F_σ -set of values of logarithmic capacity zero (**Hällström-Kametani theorem**). For the study of value distribution at general singularities, it is useful to investigate cluster sets (\rightarrow 62 Cluster Sets). In order to generalize the Nevanlinna theory to the case of general domains or Riemann surfaces, we take their exhaustions depending on a real parameter r and define the \dagger counting function and so on (\rightarrow 272 Meromorphic Functions). G. J. Hällström established the value distribution theory of meromorphic functions defined in the complementary domain D of a compact set E of logarithmic capacity zero by taking $D_r = \{z \mid v(z) < r\}$ as the exhaustion of D , where $v(z)$ denotes the Evans potential for E , i.e., $v(z)$ is the potential corresponding to a positive mass distribution μ on E of total mass 1 which tends to $+\infty$ as z tends to any point of E . Thus the number of Picard's exceptional values is not greater than $2 + \xi$, where $\xi = \limsup_{r \rightarrow \infty} F(r)/T(r)$ with $-n(r)$ = the Euler characteristic of D_r and $F(r) = \int_0^r n(r) dr$ (Hällström-Tsuji theorem). J. Tamura, L. Sario, and others studied the value distributions of meromorphic functions defined on Riemann surfaces. Sario succeeded in extending the Nevanlinna theory to analytic mappings of a Riemann surface \mathfrak{R} into another Riemann surface \mathfrak{S} by introducing a suitable metric in \mathfrak{S} to define the \dagger proximity function.

In the Nevanlinna theory on general domains, we must sometimes impose conditions that the functions must satisfy in order to obtain a concrete conclusion, for instance, the condition that ξ be finite in the Hällström-Tsuji theorem. But it is also important to determine the domains where some result can

be obtained without imposing any additional conditions on the functions. The Hällström-Kametani theorem is an example. Although the set of exceptional values in this theorem cannot be replaced generally by a smaller set than an F_σ -set of logarithmic capacity zero, we have the following theorem: Let E be a \dagger Cantor set with successive ratios $\xi_n = 2l_n/l_{n-1}$, where l_n denotes the length of the segments that remain after repeating n times the process of deleting an open segment from the middle of another segment. Then any single-valued meromorphic function with E as the set of singularities has at most 3 Picard's exceptional values if $\lim_{n \rightarrow \infty} \xi_n = 0$ and at most 2 if $\xi_{n+1} = o(\xi_n^2)$ (L. Carleson, K. Matsumoto). By weakening the conditions on E , one can get several improvements of this result.

References

- [1] K. Noshiro, Modern theory of functions (in Japanese), Iwanami, 1954.
- [2] A. Dinghas, Vorlesungen über Funktionentheorie, Springer, 1961.
- [3] W. K. Hayman, Meromorphic functions, Clarendon Press, 1964.
- [4] H. Cartan, Sur les zéros des combinaisons linéaires de p fonctions holomorphes données, *Mathematica*, 7 (1933), 5–31.
- [5] L. V. Ahlfors, The theory of meromorphic curves, *Acta Soc. Sci. Fenn.*, ser. A, 3 (1941), 3–31.
- [6] H. and J. Weyl, Meromorphic functions and analytic curves, *Ann. Math. Studies* 12, Princeton Univ. Press, 1943.
- [7] M. Cowen and P. Griffiths, Holomorphic curves and metrics of negative curvature, *J. Analyse Math.*, 29 (1976), 93–153.
- [8] L. Sario and K. Noshiro, Value distribution theory, Van Nostrand, 1966.
- [9] K. Matsumoto, Existence of perfect Picard sets, *Nagoya Math. J.*, 27 (1966), 213–222.

125 (XII.7) Distributions and Hyperfunctions

A. History

The advancement of analysis, particularly in the field of partial differential equations and harmonic analysis, necessitated the generalization of the notion of functions and derivatives. For instance, "functions" such as Dirac's \dagger delta function and \dagger Heaviside's function were used by physicists and engineering scientists

Distributions and Hyperfunctions

even though the former is not a function and the latter is not a differentiable function in the classical sense. The finite parts of divergent integrals, used by J. Hadamard to investigate the fundamental solutions of wave equations (1932), and the Riemann-Liouville integrals due to M. Riesz (1938) were the notions that eventually led to the theory of generalized functions. The rudiments of the idea of distribution, however, can also be found in other earlier works. S. Bochner (1932) and T. Carleman (1944) discussed the Fourier transforms of locally integrable functions on the reals with growth as large as a polynomial. S. L. Sobolev introduced the notion of generalized derivative and also of generalized solution of differential equations by means of integration by parts in studying the Cauchy problem for hyperbolic equations (1936); J. Leray (1934), K. O. Friedrichs (1939), and C. B. Morrey, Jr. (1940) also discussed generalized derivatives. On the other hand, L. Fantappiè (1943) investigated analytic functionals that are elements of the dual of the space of analytic functions and applied them to the theory of partial differential equations. Based on a systematic generalization of these investigations, L. Schwartz [1] established the theory of distributions (1945), which not only provided a mathematical foundation for a number of formal methods that had been used in mathematical physics but also gave new and powerful tools for the theories of differential equations (L. Hörmander [2]) and \dagger Fourier transforms. Furthermore, it has been applied to \dagger representation theory of locally compact groups, the theory of probability, and the theory of manifolds (G. de Rham [3]). As will be seen in Section B, distributions are defined as continuous \dagger functionals on a certain function space, and it is essential to select a function space appropriate to the problems concerned. For this reason, I. M. Gel'fand and G. E. Shilov defined various classes of generalized functions [4] as a natural extension of Schwartz's theory. In this direction there are also various classes of ultradistributions introduced by C. Roumieu [5] and A. Beurling.

Another but completely different approach was given by M. Sato [6, 15] in the form of the theory of hyperfunctions (1958). Intuitively a hyperfunction is the sum of (ideal) boundary values of holomorphic functions at the real axis. We obtain in this way an ultimate class of localizable generalized functions. Sato employed the relative (or local) cohomology theory to define the "boundary values" and to prove their localizing property. Such an algebraic approach to generalized functions led naturally to an algebraic treatment of systems of linear partial differential operators [7], called **algebraic analysis** by comparison with

algebraic geometry. Independently, Hörmander established a similar theory for distributions using Fourier analysis (\rightarrow 274 Micro-local Analysis).

B. Definition of Distributions

Let $\varphi(x)$ be a complex-valued function of $x = (x_1, \dots, x_n)$ defined on an open set Ω in the n -dimensional Euclidean space \mathbf{R}^n . By the **support** (or **carrier**) of φ , denoted by $\text{supp } \varphi$, we mean the \dagger closure of $\{x \mid \varphi(x) \neq 0\}$ in Ω . For multi-indices p , i.e., n -tuples $p = (p_1, \dots, p_n)$ of nonnegative integers, we set $|p| = p_1 + \dots + p_n$. For a function $\varphi(x)$ of class $C^{|p|}$, we write

$$D^p \varphi = \frac{\partial^{|p|} \varphi(x)}{\partial x_1^{p_1} \dots \partial x_n^{p_n}}. \quad (1)$$

In particular, $D^{(0, \dots, 0)} \varphi = \varphi$. To indicate the variables x we shall adopt the notation D_x^p .

$\mathcal{D}(\Omega)$ denotes the set of all complex-valued functions of class C^∞ defined on Ω with \dagger compact support, which is a \dagger linear space under the usual addition and scalar multiplication in function spaces. A sequence $\{\varphi_m\}$ in $\mathcal{D}(\Omega)$ is said to converge to 0 (the function identically equal to zero) as $m \rightarrow \infty$, denoted by $\varphi_m \Rightarrow 0$, if there exists a compact set E in Ω such that E contains $\text{supp } \varphi_m$ for every m , and for every p , $\{D^p \varphi_m\}$ converges uniformly to 0 as $m \rightarrow \infty$.

We sometimes abbreviate $\mathcal{D}(\Omega)$ either as \mathcal{D} or, when we want to indicate the variables x , as \mathcal{D}_x .

A complex-valued \dagger linear functional T defined on $\mathcal{D}(\Omega)$ is called a **distribution on Ω** if it is continuous on $\mathcal{D}(\Omega)$, i.e., $\varphi_m \Rightarrow 0$ implies $T(\varphi_m) \rightarrow 0$. The set of all distributions on Ω is denoted by $\mathcal{D}'(\Omega)$ (or \mathcal{D}'). For distributions S and T , the sum $S + T$ and scalar multiple αT are defined by $(S + T)(\varphi) = S(\varphi) + T(\varphi)$ and $(\alpha T)(\varphi) = \alpha T(\varphi)$, respectively, which are also distributions. Hence $\mathcal{D}'(\Omega)$ is a \dagger linear space.

C. Examples of Distributions

(1) Let $f(x)$ be a \dagger measurable and \dagger locally integrable function on Ω . Then a distribution T_f is defined by $T_f(\varphi) = \int \varphi(x) f(x) dx$. Here dx is the \dagger Lebesgue measure and the domain of integration is Ω (in fact, $\text{supp } \varphi$). If $T_f = T_g$, then $f(x) = g(x)$ \dagger almost everywhere. Thus we can identify the distribution T_f with the corresponding function f , and sometimes T_f will be denoted simply by f .

(2) Let μ be a \dagger Radon measure on Ω , i.e., a complex-valued \dagger regular completely additive set function on the \dagger Borel sets in Ω . Then a distribution T_μ is defined by $T_\mu(\varphi) = \int \varphi(x) \mu(dx)$. Example (1) is a special case where

$\mu(dx) = f(x)dx$. Two Radon measures μ and ν coincide if $T_\mu = T_\nu$. If the measure is concentrated at the origin, then $T_\mu(\varphi) = c\varphi(0)$, denoted by $c\delta$, and δ is called **Dirac's distribution**.

Sometimes δ is denoted by δ_x , $\delta_{(x)}$, or $\delta(x)$ to indicate that it operates on functions of x . A distribution $T \in \mathcal{D}'(\Omega)$ is a measure if and only if $T(\varphi_m) \rightarrow 0$ whenever the supports $\text{supp } \varphi_m$ of a sequence $\varphi_m \in \mathcal{D}(\Omega)$ are in a fixed compact set in Ω and φ_m converges uniformly to zero. A distribution T is called a **positive distribution** if $T(\varphi) \geq 0$ for any $\varphi \in \mathcal{D}(\Omega)$ that has non-negative values at all points x in Ω . Every positive distribution is equal to a T_μ corresponding to a positive measure μ .

(3) For given p the distribution $\delta^{(p)}$ is defined by $\delta^{(p)}(\varphi) = (-1)^{|p|} D^p \varphi(0) \cdot \delta^{(0, \dots, 0)} = \delta$.

(4) Let $g(x)$ be a function defined but not integrable on an interval (a, b) , and assume that for any positive number ε it is integrable on $(a + \varepsilon, b)$. Moreover, assume that

$$g(x) = \sum_{v=1}^n A_v (x-a)^{-\lambda_v} + h(x),$$

where $\text{Re } \lambda_v > 1$, λ_v is not an integer, and $h(x)$ is integrable on (a, b) . Then

$$\int_{a+\varepsilon}^b g(x) dx - \sum A_v (\lambda_v - 1)^{-1} \varepsilon^{1-\lambda_v} = F(\varepsilon)$$

tends to a finite value as $\varepsilon \rightarrow 0$. This limit is called the **finite part** (in French *partie finie*) of the integral $\int_a^b g(x) dx$, denoted by $\text{Pf} \int_a^b g(x) dx$:

$$\begin{aligned} \text{Pf} \int_a^b g(x) dx \\ = - \sum_v \frac{A_v}{\lambda_v - 1} (b-a)^{1-\lambda_v} + \int_a^b h(x) dx. \end{aligned}$$

In the same way, for every $\varphi \in \mathcal{D}(\mathbf{R}^1)$, $T(\varphi) = \text{Pf} \int_a^b g(x) \varphi(x) dx$ is defined, and T is a distribution which will be denoted by $\text{Pf} g$, and which is frequently called a **pseudofunction**. This notion is extended to the n -dimensional case and used to express the fundamental solutions (\rightarrow Section EE; Appendix A, Table 15.V) of \dagger hyperbolic partial differential equations.

D. Localization of Distributions

Let Ω' be an open subset of Ω . Every function $\varphi \in \mathcal{D}(\Omega')$ can be extended to a function $\bar{\varphi} \in \mathcal{D}(\Omega)$ by setting $\bar{\varphi}(x) = 0$ for $x \notin \Omega'$. Thus if $T \in \mathcal{D}'(\Omega)$, then a distribution $S \in \mathcal{D}'(\Omega')$ is defined by $S(\varphi) = T(\bar{\varphi})$ for every $\varphi \in \mathcal{D}(\Omega')$, and S is called the **restriction** of T to Ω' . Two distributions T and S are said to be equal in Ω' if their restrictions to Ω' are equal. If every point in Ω has a neighborhood where $T = S$, then $T = S$. In this sense a distribution is determined completely by its local data, although the

notion of pointwise value, having exact meaning for functions, has no meaning for distributions. Moreover, we can construct a distribution with given local data in the following way: Suppose that an \dagger open covering $\{\Omega_j\}$ of Ω and a set of distributions $T_j \in \mathcal{D}'(\Omega_j)$ are given and that T_j and T_k are equal in $\Omega_j \cap \Omega_k$ for any j and k . Then there exists a unique distribution $T \in \mathcal{D}'(\Omega)$ such that $T = T_j$ in Ω_j for each j . (Define $T(\varphi) = \sum T_j(\alpha_j \varphi)$ for a \dagger partition of unity $\{\alpha_j\}$ subordinate to $\{\Omega_j\}$). Namely, the distributions \mathcal{D}' form a \dagger sheaf of linear spaces over \mathbf{R}^n . It is not \dagger flabby but is soft, i.e., for any distribution T defined on a neighborhood of a closed set F in \mathbf{R}^n , there is a distribution on \mathbf{R}^n that coincides with T on a neighborhood of F .

For each distribution $T \in \mathcal{D}'(\Omega)$ there is a largest open subset Ω' of Ω on which T vanishes. Its complement is called the **support** (or **carrier**) of T and is denoted by $\text{supp } T$. The support of the distribution given in example (1) coincides with that of the function f . The support of $\delta^{(p)}$ is the origin.

E. Derivatives of Distributions

In example (1) in Section C, if $f(x)$ is a function of class C^k , then by integration by parts $T_{D^p f}(\varphi) = (-1)^{|p|} T_f(D^p \varphi)$ for $|p| \leq k$. The right-hand side defines a distribution even if f is not differentiable. In view of this example, we define **derivatives** $D^p T$ of any distribution T by

$$(D^p T)(\varphi) = (-1)^{|p|} T(D^p \varphi), \quad \varphi \in \mathcal{D}_x. \quad (2)$$

Any distribution is infinitely differentiable. Any locally integrable function is infinitely differentiable in the sense of distributions, and its derivatives $D^p T_f$ are called **distribution derivatives** (or **generalized derivatives** or **weak derivatives**).

For example: (1) $D^p \delta = \delta^{(p)}$; (2) (1-dimensional case) $dx_+/dx = 1$, $d1/dx = \delta$, where $x_+ = \max(x, 0)$ and $1(x)$ is **Heaviside's function**, which is equal to 1 for $x \geq 0$ and to 0 for $x < 0$.

F. The Operation of Linear Differential Operators on Distributions

Let T be a distribution and $\alpha(x)$ a C^∞ -function on Ω . Then we can define the product αT by $(\alpha T)(\varphi) = T(\alpha \varphi)$, where the right-hand side is a continuous linear functional on $\mathcal{D}(\Omega)$ since the multiplication $\varphi \mapsto \alpha \varphi$ is continuous in $\mathcal{D}(\Omega)$. We define the **dual operator** (or **conjugate operator**) $P'(x, D)$ of a linear differential operator $P(x, D) = \sum a_p(x) D^p$ by

$$P'(x, D)\varphi = \sum (-1)^{|p|} D^p(a_p(x)\varphi(x)), \quad (3)$$

where the $a_p(x)$ are C^∞ -functions on Ω . Com-

binning differentiation of distributions, multiplication by functions, and addition, we can apply partial differential operators to distributions, and we have $(P(x, D)T)(\varphi) = T(P'(x, D)\varphi)$.

Linear differential operators commute with restriction mappings. In particular, we have $\text{supp } P(x, D)T \subset \text{supp } T$. In other words, linear differential operators are \dagger sheaf homomorphisms on the sheaf \mathcal{D}' of distributions over Ω . On the other hand, every continuous sheaf homomorphism on \mathcal{D}' is a linear differential operator whose order is finite on each compact set in Ω . Even when no continuity is assumed, a sheaf homomorphism on \mathcal{D}' is a linear differential operator except on a discrete subset of Ω (J. Peetre).

G. The Topology of \mathcal{D} and \mathcal{D}'

Let $\emptyset = K_1 \subset K_2 \subset \dots$ be a sequence of compact sets in Ω that exhausts Ω . For every nondecreasing sequence of positive numbers $\{a_j\} = \{a_0, a_1, \dots\}$ and nondecreasing sequence of nonnegative integers $\{k_j\} = \{k_0, k_1, \dots\}$ we set

$$\rho_{\{a_j\}, \{k_j\}}(\varphi) = \sup_{j \geq 0} \sup_{|p| \leq k_j} \sup_{x \notin K_j} |D^p \varphi(x)| \quad (4)$$

for $\varphi \in \mathcal{D}(\Omega)$. We define a \dagger locally convex topology of the space $\mathcal{D}(\Omega)$ by taking the totality of the \dagger seminorms $\rho_{\{a_j\}, \{k_j\}}$ as a fundamental system of continuous seminorms. Then $\mathcal{D}(\Omega)$ is a \dagger nuclear \dagger (LF)-space, and the convergence $\varphi_m \rightarrow 0$ of a sequence φ_m is identical to the convergence $\varphi_m \rightarrow 0$ in this topology. A subset $B \subset \mathcal{D}(\Omega)$ is \dagger bounded in the topology if and only if there exist a compact set $E \subset \Omega$ such that $\text{supp } \varphi \subset E$ for every $\varphi \in B$, and positive numbers M_p for each p such that $\sup |D^p \varphi(x)| \leq M_p$ for every $\varphi \in B$.

The topology of the space $\mathcal{D}'(\Omega)$ is the \dagger strong topology of the \dagger dual of $\mathcal{D}(\Omega)$; the \dagger topology of uniform convergence on every bounded set in $\mathcal{D}(\Omega)$. Under this topology $\mathcal{D}'(\Omega)$ is a \dagger nuclear \dagger reflexive linear topological space. With respect to this topology, the linear differential operators of Section F are continuous in \mathcal{D}' .

By virtue of the following convergence theorems, various limiting processes concerning distributions can be treated easily. If the limit $\lim T_j(\varphi) = T(\varphi)$ exists for every $\varphi \in \mathcal{D}$, where $\{T_j\}$ is a sequence of distributions, then $T \in \mathcal{D}'$ and $\{T_j\}$ is convergent to the distribution T (**convergence theorem**). Moreover, for any p , $D^p T_j$ is convergent to $D^p T$ (**theorem of termwise differentiation**). Any bounded set in \mathcal{D} or \mathcal{D}' is totally bounded. Thus weak convergence of a sequence $\{T_j\}$ implies strong convergence (convergence in the topology of the space \mathcal{D}') (**strong convergence theorem**).

H. Distributions Depending on a Parameter

Consider distributions T_λ depending on a parameter λ , where λ ranges over the real line, the complex plane, or more generally an open set in Euclidean space. The convergence theorem and the strong convergence theorem also hold in the case of a continuous parameter λ . Thus T_λ is **continuous (differentiable) with respect to the parameter λ** if $T_\lambda(\varphi)$ is continuous (differentiable) with respect to λ for any $\varphi \in \mathcal{D}$. If T_λ is continuous with respect to λ , then $D_x^p T_\lambda$ is also continuous with respect to λ . If T_λ is differentiable with respect to a real variable λ , then $D_x^p T_\lambda$ is also differentiable, and $\partial D_x^p T_\lambda / \partial \lambda = D_x^p (\partial T_\lambda / \partial \lambda)$. The same facts hold for the case of several real variables. For a complex parameter λ , T_λ is **holomorphic with respect to λ** if $T_\lambda(\varphi)$ is holomorphic for any $\varphi \in \mathcal{D}$. The fundamental properties of holomorphic functions also hold in this case.

If T_λ is defined and continuous in an interval $[a, b]$, then the **integral $T = \int T_\lambda d\lambda$ with respect to λ** exists in \mathcal{D}' and we have

$$T(\varphi) = \int_a^b T_\lambda(\varphi) d\lambda, \quad \varphi \in \mathcal{D}.$$

I. Distributions with Compact Support

We denote by $\mathcal{E}(\Omega)$ (abbreviated to \mathcal{E}) the space of all complex valued C^∞ -functions on Ω . $\mathcal{E}(\Omega)$ is a nuclear \dagger Fréchet space with the locally convex topology defined by the seminorms

$$\rho_{p, K}(\varphi) = \sup_{x \in K} |D^p \varphi(x)| \quad (5)$$

as p ranges over all multi-indices and K ranges over all compact sets in Ω .

We denote by $\mathcal{E}'(\Omega)$ (abbreviated to \mathcal{E}') the \dagger strong dual of $\mathcal{E}(\Omega)$, i.e., the set of all continuous linear functionals on $\mathcal{E}(\Omega)$ equipped with the topology of uniform convergence on bounded sets in $\mathcal{E}(\Omega)$. $\mathcal{E}'(\Omega)$ is a nuclear \dagger (DF)-space. If $T \in \mathcal{E}'(\Omega)$, then its restriction to $\mathcal{D}(\Omega)$ is a distribution with compact support. Conversely if T is a distribution with compact support in Ω , then choosing $\alpha \in \mathcal{D}(\Omega)$ which is identically equal to 1 on a neighborhood of the support of T , we define a linear functional S on $\mathcal{E}(\Omega)$ by $S(\varphi) = T(\alpha\varphi)$. Then $S \in \mathcal{E}'(\Omega)$, and S is independent of the choice of α . In this sense, we can consider that S is the extension of T to $\mathcal{E}(\Omega)$ and identify S and T ; thus $\mathcal{E}'(\Omega)$ coincides with the set of all distributions with compact support in Ω .

J. Structure Theorems for Distributions

A distribution $T \in \mathcal{D}'(\Omega)$ is said to be of order at most k if $|T(\varphi)| \leq \rho_{\{a_j\}, \{k_j\}}(\varphi)$ for $\{k_j\} = \{k, k, k, \dots\}$

and for some $\{a\}$. A distribution of order 0 is a measure. Every **distribution of finite order** can be represented as a finite linear combination of derivatives (in the sense of distributions) of measures or locally integrable functions. The restriction of any distribution to a relatively compact open subset Ω' can be represented in this way because it is of finite order. Therefore the distributions form the smallest class of generalized functions that contains all locally integrable functions, is stable under differentiation, and forms a sheaf. Furthermore, for any distribution T there exists a sequence $\{\varphi_j\} \subset \mathcal{D}$ that converges to T in the distribution sense.

A closed set F is said to be **regular** if for every point a in F we have a neighborhood U of a and constants ω and $1 \geq \alpha > 0$ such that every pair of points x and y in $F \cap U$ is connected by a curve contained in F with length less than or equal to $\omega|x - y|^\alpha$. If the support of a distribution T is contained in a regular closed set F , then

$$T = \sum D^{p_j} T_{\mu_j}$$

(the expression is not necessarily unique, and the sum is locally finite), where the μ_j are complex-valued measures with support contained in F . In particular, if the support of a distribution T contains only one point a , then T can be represented uniquely as a finite linear combination

$$T = \sum_{|p| \leq m} \alpha_p D^p \delta_{(x-a)}$$

of derivatives of the distribution $\delta_{(x-a)}$ defined by $\delta_{(x-a)}(\varphi) = \varphi(a)$.

K. Tensor Products of Distributions

Let $\Omega_x \subset \mathbf{R}^n$ and $\Omega_y \subset \mathbf{R}^m$ be open sets. If $T \in \mathcal{D}'(\Omega_x)$ and $S \in \mathcal{D}'(\Omega_y)$, then there is a unique distribution $T \otimes S \in \mathcal{D}'(\Omega_x \times \Omega_y)$, called the **tensor product** (or **direct product**), such that $T \otimes S(\varphi(x)\psi(y)) = T(\varphi)S(\psi)$ for any $\varphi \in \mathcal{D}(\Omega_x)$ and $\psi \in \mathcal{D}(\Omega_y)$. More directly we have Fubini's theorem: $T \otimes S(\varphi(x, y)) = T_x(S_y(\varphi(x, y))) = S_y(T_x(\varphi(x, y)))$ for any $\varphi \in \mathcal{D}(\Omega_x \times \Omega_y)$. If F and G are linear spaces of distributions on Ω_x and Ω_y , respectively, then the \dagger tensor product $F \otimes G$ is identified with the linear combinations of tensor products $T \otimes S$ of $T \in F$ and $S \in G$, which form a linear space of distributions on $\Omega_x \times \Omega_y$. When F and G have \dagger locally convex topologies, the \dagger completed tensor products $F \hat{\otimes} G$ and $F \hat{\otimes} G$ are also usually identified with subspaces of $\mathcal{D}'(\Omega_x \times \Omega_y)$. For example, we have $\mathcal{D}'(\Omega_x) \hat{\otimes} \mathcal{D}'(\Omega_y) = \mathcal{D}'(\Omega_x \times \Omega_y)$, $\mathcal{E}'(\Omega_x) \hat{\otimes} \mathcal{E}'(\Omega_y) = \mathcal{E}'(\Omega_x \times \Omega_y)$ and $\mathcal{E}(\Omega_x) \hat{\otimes} \mathcal{E}(\Omega_y) = \mathcal{E}(\Omega_x \times \Omega_y)$ (including the topologies). Similarly, we have $\mathcal{D}(\Omega_x) \hat{\otimes} \mathcal{D}(\Omega_y) = \mathcal{D}(\Omega_x \times \Omega_y)$,

but the left-hand side has a strictly weaker topology than the right-hand side. Spaces of distributions with parameters (\rightarrow Section H) are often identified with completed tensor products of spaces of distributions [8–11].

L. The Kernel Theorem

Let Ω_x and Ω_y be as above. Every distribution K on $\Omega_x \times \Omega_y$ induces the continuous linear mapping $L_K: \mathcal{D}(\Omega_x) \rightarrow \mathcal{D}'(\Omega_y)$ defined by $(L_K(\varphi(x)))(\psi(y)) = K(\varphi(x)\psi(y))$, $\psi \in \mathcal{D}(\Omega_y)$. Conversely every continuous linear mapping $L: \mathcal{D}(\Omega_x) \rightarrow \mathcal{D}'(\Omega_y)$ is equal to L_K for a $K \in \mathcal{D}'(\Omega_x \times \Omega_y)$ and the correspondence $K \mapsto L_K$ is a topological isomorphism of the locally convex space $\mathcal{D}'(\Omega_x \times \Omega_y)$ onto $L(\mathcal{D}(\Omega_x), \mathcal{D}'(\Omega_y))$ equipped with the topology of uniform convergence on the bounded sets (L. Schwartz's kernel theorem, [8–11]). K is called the **kernel (distribution)** of the mapping L_K .

L_K is a continuous linear mapping $\mathcal{D}(\Omega_x) \rightarrow \mathcal{E}(\Omega_y)$ (resp. $\mathcal{D}(\Omega_x) \rightarrow \mathcal{E}'(\Omega_y)$) if and only if $K \in \mathcal{D}'(\Omega_x) \hat{\otimes} \mathcal{E}(\Omega_y)$ (resp. $\mathcal{D}'(\Omega_x) \hat{\otimes} \mathcal{E}'(\Omega_y)$), in which case K is said to be **regular** (resp. **compact**) in y . L_K can be extended to a continuous linear mapping $\mathcal{E}'(\Omega_x) \rightarrow \mathcal{D}'(\Omega_y)$ (resp. $\mathcal{E}(\Omega_x) \rightarrow \mathcal{D}'(\Omega_y)$) if and only if $K \in \mathcal{E}'(\Omega_x) \hat{\otimes} \mathcal{D}'(\Omega_y)$ (resp. $\mathcal{E}'(\Omega_x) \hat{\otimes} \mathcal{D}'(\Omega_y)$), in which case K is said to be **regular** (resp. **compact**) in x . K is said to be **regular** (resp. **compact**) if it is regular (resp. compact) both in x and y . L_K can be extended to a continuous linear mapping $\mathcal{E}'(\Omega_x) \rightarrow \mathcal{E}(\Omega_y)$ (resp. $\mathcal{E}(\Omega_x) \rightarrow \mathcal{E}'(\Omega_y)$) if and only if $K \in \mathcal{E}'(\Omega_x \times \Omega_y)$ (resp. $\mathcal{E}'(\Omega_x \times \Omega_y)$). Then K is said to be **regularizing** (resp. **compactifying**) [8, 10].

M. Convolution

For distributions S and T on \mathbf{R}^n , assume that either S or T has a compact support or more generally that $\text{supp } T \cap (\{x\} - \text{supp } S)$ is locally uniformly bounded with respect to x . Then the correspondence

$$\varphi \in \mathcal{D} \rightarrow S_{(x)} T_{(y)}(\varphi(x+y))$$

defines a distribution called the **convolution** of S and T and is denoted by $S * T$. In particular, $T_f * T_g = T_{f * g}$, where $f * g$ is the \dagger convolution of functions f and g , and T_f , T_g , and $T_{f * g}$ denote the distributions corresponding to f , g , and $f * g$, respectively. For example: $T * \delta = \delta * T = T$, $D^p T * S = T * (D^p S) = D^p (T * S)$.

Thus a solution of the partial differential equation $P(D)T = S$ with constant coefficients is given by the convolution $S * E$ of S with a \dagger fundamental solution E of $P(D)$, whenever the convolution exists.

If $T \in \mathcal{D}'$ and $\varphi \in \mathcal{D}$, the convolution $T * \varphi$ is equal to a function f of class C^∞ (or the

distribution corresponding to f), and $f(y) = T_{(x)}(\varphi(y-x))$. f is called the **regularization** of T . For distributions S and T , assume that $|f(x)g(y-x)|$ is integrable on \mathbf{R}^n for any regularization $f = S * \varphi$ and $g = T * \psi$. Then there exists a unique distribution V such that

$$V * (\varphi * \psi) = \int f(x)g(y-x)dx.$$

This distribution V is called the **generalized convolution** and is denoted by $S * T$ (C. Chevalley).

N. Tempered Distributions \mathcal{S}'

We denote by $\mathcal{S}(\mathbf{R}^n)$ (abbreviated to \mathcal{S}) the space of all † rapidly decreasing functions of class C^∞ on \mathbf{R}^n . \mathcal{S} is a nuclear Fréchet space with the locally convex topology defined by the seminorms

$$\rho_{p,q}(\varphi) = \sup_x |x^p D^q \varphi(x)|, \quad (6)$$

as p and q range over all multi-indices, where $x^p = x_1^{p_1} \dots x_n^{p_n}$. We denote by $\mathcal{S}'(\mathbf{R}^n)$ (abbreviated to \mathcal{S}') the strong dual of $\mathcal{S}(\mathbf{R}^n)$. Its elements are called **tempered distributions** (or **slowly increasing distributions**) and are identified with their restrictions to $\mathcal{D}(\mathbf{R}^n)$. A distribution $T \in \mathcal{D}'$ can be continuously extended to \mathcal{S} if and only if any regularization $f = T * \varphi$ of T is a slowly increasing continuous function (i.e., there is a polynomial $P(x)$ such that $|f(x)| \leq |P(x)|$). $\mathcal{S}'(\mathbf{R}^n)$ is a nuclear (DF)-space. Let $1 < p < \infty$, and let q be the † conjugate exponent. The strong dual of the † function space $\mathcal{D}_{L_q}(\mathbf{R}^n)$ is denoted by $\mathcal{D}'_{L_p}(\mathbf{R}^n)$. Since \mathcal{S} is dense in \mathcal{D}_{L_q} , \mathcal{D}'_{L_p} is identified with a linear subspace of \mathcal{S}' . The strong dual of $\mathcal{D}'_{L_p}(\mathbf{R}^n)$ is equal to $\mathcal{D}_{L_q}(\mathbf{R}^n)$. Let $\mathcal{B}(\mathbf{R}^n)$ be the closed linear subspace of $\mathcal{B}(\mathbf{R}^n)$ consisting of all functions φ such that $D^p \varphi \in C_\infty(\mathbf{R}^n)$ for all p (— 168 Function Spaces). It is also the closure of \mathcal{S} in \mathcal{B} . The strong dual of \mathcal{B} is denoted by $\mathcal{B}'_{L_1}(\mathbf{R}^n)$. Its elements are called **integrable distributions** because the strong dual of $\mathcal{D}'_{L_1}(\mathbf{R}^n)$ is equal to $\mathcal{B}(\mathbf{R}^n)$ so that the integral $T(1)$ has a meaning. Here 1 is the function identically equal to 1.

† Sobolev spaces $W_p^l(\mathbf{R}^n)$ and † Besov spaces $B_{p,q}^s(\mathbf{R}^n)$ are also linear subspaces of $\mathcal{S}'(\mathbf{R}^n)$.

O. Fourier Transforms

The Fourier transform

$$\mathcal{F}\varphi(x) = (\sqrt{2\pi})^{-n} \int \varphi(\xi) \exp(-ix\xi) d\xi \quad (7)$$

is an isomorphism of \mathcal{S}_ξ onto \mathcal{S}_x , where $x\xi = x_1\xi_1 + \dots + x_n\xi_n$. The Fourier transform

$\mathcal{F}T \in \mathcal{S}'_\xi$ of $T \in \mathcal{S}'_x$ is defined by

$$(\mathcal{F}T)(\varphi) = T(\mathcal{F}\varphi), \quad \varphi \in \mathcal{S}. \quad (8)$$

The **inverse Fourier transform** is defined similarly, except that $-i$ is replaced by i . For example: (1) $\mathcal{F}1 = (\sqrt{2\pi})^n \delta$, where 1 is the distribution corresponding to the function 1. (2) $\mathcal{F}(D^p T) = i^{|p|} x^p \mathcal{F}T$.

A function φ of class C^∞ is called a **slowly increasing C^∞ -function** if φ and its derivatives of any order are slowly increasing continuous functions. The space \mathcal{O}_M is the set of all such functions. Its image under the Fourier transform $\mathcal{F}(\mathcal{O}_M)$, denoted by \mathcal{O}'_C , coincides with the set of all distributions T such that any regularization $f = T * \varphi$ is a rapidly decreasing C^∞ -function. A member of the space \mathcal{O}'_C is called a **rapidly decreasing distribution**. If $\alpha \in \mathcal{O}_M$ ($\beta \in \mathcal{O}'_C$) and $T \in \mathcal{S}'$, then $\alpha T \in \mathcal{S}'$ and $\mathcal{F}(\alpha T) = (\sqrt{2\pi})^{-n} \mathcal{F}\alpha * \mathcal{F}T$ ($\beta * T \in \mathcal{S}'$ and $\mathcal{F}(\beta * T) = (\sqrt{2\pi})^n \mathcal{F}\beta \mathcal{F}T$). If T is a distribution with compact support in \mathbf{R}^n , then its Fourier transform is the entire function $f(\zeta) = (\sqrt{2\pi})^{-n} T(\exp(-ix\xi))$, where $\zeta = \xi + i\eta \in \mathbf{C}^n$. For a convex compact set K in \mathbf{R}^n , define its **supporting function** H_K by $H_K(\eta) = \sup_{x \in K} x\eta$.

Paley-Wiener theorem. An entire function $f(\zeta)$ on \mathbf{C}^n is the Fourier transform of a distribution (resp. a function of class C^∞) with support in K if and only if (i) for any $\varepsilon > 0$ there is a constant C_ε such that $|f(\zeta)| \leq C_\varepsilon \exp(H_K(\eta) + \varepsilon|\zeta|)$ on \mathbf{C}^n , and (ii) there are constants C and N such that $|f(\xi)| \leq C(1+|\xi|)^N$ (resp. for any N there is a constant C such that $|f(\xi)| \leq C(1+|\xi|)^{-N}$) on \mathbf{R}^n .

P. Fourier Series and Distributions on Tori

The n -dimensional † torus \mathbf{T}^n is the quotient space of \mathbf{R}^n with respect to the equivalence relation $x_j \equiv y_j \pmod{\mathbf{Z}}$ ($j = 1, \dots, n$). The space $\mathcal{D}'(\mathbf{T}^n)$ of distributions on \mathbf{T}^n is defined to be the strong dual of $\mathcal{D}(\mathbf{T}^n) = \mathcal{E}(\mathbf{T}^n)$. The volume element dx of \mathbf{T}^n is defined from that of \mathbf{R}^n . Thus, for an integrable function f , we can define a distribution T_f by

$$T_f(\varphi) = \int_{\mathbf{T}^n} f(x)\varphi(x)dx, \quad \varphi \in \mathcal{D}(\mathbf{T}^n).$$

Consider the family of functions $f_p(x) = \exp(2\pi i p x)$, where p ranges over all n -tuples of integers. Then any distribution T on \mathbf{T}^n has the **Fourier series** expansion $T = \sum c_p f_p(x)$, where the Fourier coefficients $c_p = T(f_{-p})$ are slowly increasing, i.e., $|c_p| \leq C(1+|p|^2)^k$ for some k and C . Conversely if $\{c_p\}$ is a slowly increasing sequence, then $\sum c_p f_p$ converges to a distribution T on \mathbf{T}^n .

Q. Substitution

Let $f = (f_1, \dots, f_m)$ be a C^∞ -mapping from an open set Ω in \mathbf{R}^n into an open set $\tilde{\Omega}$ in \mathbf{R}^m , and assume that the t rank of the Jacobian matrix $(\partial f_i / \partial x_j)$ ($i = 1, \dots, m; j = 1, \dots, n$) is equal to m in a neighborhood of the t inverse image E of the support of a distribution $S = S_{(y)} \in \mathcal{D}'(\tilde{\Omega})$. In a neighborhood U of E , we can choose $u_1 = g_1(x), \dots, u_{n-m} = g_{n-m}(x)$ so that the transformation $(y, u) = (f(x), g(x))$ has the inverse transformation $x = \psi(y, u)$ of class C^∞ . If the support of $\varphi \in \mathcal{D}(\Omega)$ is contained in U , then defining $\tilde{\varphi}$ by

$$\tilde{\varphi}(y) = \int_{\mathbf{R}^{n-m}} \varphi(\psi(y, u)) J(y, u) du,$$

we have $\tilde{\varphi} \in \mathcal{D}(\tilde{\Omega})$, where $J(y, u)$ is the absolute value of the t Jacobian of $\psi(y, u)$, and $\tilde{\varphi}$ is independent of the choice of g_1, \dots, g_{n-m} . Now we define $T(\varphi) = S(\tilde{\varphi})$ if $\text{supp } \varphi \subset U$, and $T(\varphi) = 0$ if $\text{supp } \varphi$ does not intersect E . Since a distribution can be determined by its local behavior, we have a distribution $T \in \mathcal{D}'(\Omega)$ with support E , which is denoted by $T = S \circ f = S(f) = S_{f(x)}$ and is called the **substituted distribution** of $S_{(y)}$ by $y = f(x)$. It is also called the **pullback** of S by f and is denoted by f^*S . The chain rule for derivatives of composites

$$\frac{\partial}{\partial x_j} (S \circ f) = \sum_{i=1}^m \frac{\partial f_i}{\partial x_j} \left(\frac{\partial S}{\partial y_i} \circ f \right) \quad (9)$$

also holds.

For example: (1) If $S = \delta_{(y)}^{(q)}$ ($y \in \mathbf{R}^m$), then E is the surface $f_1(x) = f_2(x) = \dots = f_m(x) = 0$. Assuming that (f_1, \dots, f_m) satisfies the condition in the previous paragraph, we obtain the distribution $\delta^{(q)}(f_1, \dots, f_m)$ with support contained in a surface. From the fact $\delta^{(q)} = D_y^q \delta$, we can write

$$\delta^{(q)}(f_1, \dots, f_m) = \frac{\partial^{(q)} \delta(f_1, \dots, f_m)}{(\partial f_1)^{q_1} \dots (\partial f_m)^{q_m}}$$

(Gel'fand-Shilov notation). (2) For the mapping $f(x) = Ax + b$, where A is an $n \times n$ regular matrix and b is an n -vector, we can define the substituted distribution of S by $f = f(x)$ for any $S \in \mathcal{D}'(\mathbf{R}^n)$. For instance, $\delta_{(x-b)}(\varphi) = \varphi(b)$, $\delta_{(x^2-c^2)} = (2c)^{-1}(\delta_{(x-c)} + \delta_{(x+c)})$ ($c > 0, x \in \mathbf{R}^1$).

R. Distributions and Currents on a Differentiable Manifold

Let M be an n -dimensional t differentiable manifold of class C^∞ . Let $\{(U_\alpha, \psi_\alpha)\}$ be an t atlas. Then a **distribution** T on M is defined to be a collection of distributions T_α on $\psi_\alpha(U_\alpha) \subset \mathbf{R}^n$ such that T_β on $\psi_\beta(U_\alpha \cap U_\beta)$ is equal to $T_\alpha \circ (\psi_\alpha \circ \psi_\beta^{-1})$ for any α and β . Since the distributions obey the same transformation law as

functions under coordinate transformation, this definition has an invariant meaning, and every locally integrable function on M is regarded as a distribution. More generally, a distribution t cross section of a t vector bundle of class C^∞ is defined in the same way. The most important is the case of the $^t p$ -fold exterior power of the t cotangent bundle. Its distribution cross sections, i.e., t exterior differential forms of order p with distribution coefficients, are called **currents** of degree p . If M is t oriented, then the space of currents of degree p on M is identified with the dual space of the locally convex space $\mathcal{D}^{(n-p)}(M)$ of all differential forms of degree $n-p$, of class C^∞ and with compact support in M . For example, if C is an $(n-p)$ -dimensional t singular chain, a current T_C of degree p is defined by $T_C(\alpha) = \int_C \alpha$. If M is not t orientable, we have to consider either the double covering \tilde{M} of M or the cross sections of the tensor product of the bundle of $(n-p)$ -covectors and the orientation bundle [1, 3]. Currents were introduced by de Rham to prove his celebrated isomorphism theorem of the t de Rham cohomology groups defined by differential forms and the t singular cohomology groups defined by singular cochains (\rightarrow 105 Differentiable Manifolds).

S. Gel'fand-Shilov Generalized Functions

Let E and F be t locally convex spaces for which a continuous linear injection $i: E \rightarrow F$ with dense range is defined. Then the dual $i': F' \rightarrow E'$ is a continuous linear injection with weak*-dense range. Therefore if we shrink the function space E , we obtain a larger space E' of continuous linear functionals. Gel'fand and Shilov [4] introduced many function spaces E , called **fundamental spaces** or **test function spaces**, and defined corresponding spaces E' of **generalized functions**. Their motivations were in applications to the theory of partial differential equations, where Fourier transforms play an essential role and Schwartz's framework of tempered distributions is too restrictive. They considered a pair of function spaces E and \tilde{E} such that the Fourier transformation $\mathcal{F}: E \rightarrow \tilde{E}$ is an isomorphism of locally convex spaces. Then the Fourier transformation defined to be the dual of \mathcal{F} gives an isomorphism $\tilde{E}' \rightarrow E'$ of spaces of generalized functions. The function spaces E and \tilde{E} are often spaces of entire functions, so that the Gel'fand-Shilov generalized functions are not necessarily localizable in \mathbf{R}^n . As typical examples of spaces E and \tilde{E} we shall mention only spaces of type S in the next section.

L. Ehrenpreis's **analytically uniform spaces**

[12] are also a class of generalized function spaces introduced from a similar point of view.

T. Spaces of Type S

Let $\alpha, \beta, A, B, \bar{A}$, and \bar{B} be n -dimensional vectors. By $A > B$ we mean $A_j > B_j$ ($j = 1, \dots, n$). We use notations $A^p = A_1^{p_1} \dots A_n^{p_n}$, $p^{\alpha} = p_1^{\alpha_1} \dots p_n^{\alpha_n}$. (i) For $\alpha \geq 0$, $A > 0$ the space $S_{\alpha, A}$ consists of all C^∞ -functions φ such that seminorms

$$\rho_{q, \bar{A}}(\varphi) = \sup_x \sup_p \frac{|x^p D^q \varphi(x)|}{\bar{A}^p p^{p\alpha}} \quad (10)$$

are finite for all $\bar{A} > A$ and q . For example, $S_{0, A}$ is the space of all functions of class C^∞ with support in $\{x \mid |x_j| \leq A_j, j = 1, \dots, n\}$. (ii) For $\beta \geq 0$, $B > 0$ the space $S^{\beta, B}$ consists of all C^∞ -functions φ such that the seminorms

$$\rho_{p, \bar{B}}(\varphi) = \sup_x \sup_q \frac{|x^p D^q \varphi(x)|}{\bar{B}^q q^{\beta p}} \quad (11)$$

are finite for all $\bar{B} > B$ and q . (iii) For $\alpha, \beta \geq 0$ and $A, B > 0$, the space $S_{\alpha, A}^{\beta, B}$ consists of all C^∞ -functions such that the seminorms

$$\rho_{A, B}(\varphi) = \sup_x \sup_p \sup_q \frac{|x^p D^q \varphi(x)|}{A^p B^q p^{p\alpha} q^{\beta p}} \quad (12)$$

are finite for all $\bar{A} > A$ and $\bar{B} > B$. The topologies for these spaces are given in terms of the seminorms (10), (11), and (12), respectively. These spaces are generically called **spaces of type S**.

$$\mathcal{F}(S_{\alpha, A}) = S^{\alpha, A} \quad (\alpha > 0),$$

$$\mathcal{F}(S^{\beta, B}) = S_{\beta, B} \quad (\beta > 0),$$

$$\mathcal{F}(S_{\alpha, A}^{\beta, B}) = S_{\beta, B}^{\alpha, A} \quad (\alpha + \beta > 1).$$

$$\mathcal{F}(S_{0, A}) = S^{0, A'},$$

$$\mathcal{F}(S^{0, B}) = S_{0, B'},$$

$$\mathcal{F}(S_{\alpha, A}^{\beta, B}) = S_{\beta, B'}^{\alpha, A'} \quad (\alpha + \beta = 1),$$

where $A' = A \exp(1/A)$ and $B' = B \exp(1/B)$.

We set $S_\alpha = \bigcup S_{\alpha, A}$, $S^\beta = \bigcup S^{\beta, B}$, and $S_\alpha^\beta = \bigcup S_{\alpha, A}^{\beta, B}$, where A and B range over all positive n -dimensional vectors. $S_\alpha^\beta \neq \{0\}$ if and only if one of the following conditions is satisfied: (i) $\alpha + \beta \geq 1$, $\alpha > 0$, $\beta > 0$; (ii) $\alpha > 1$, $\beta = 0$; (iii) $\alpha = 0$, $\beta > 1$. In such a case the space S_α^β of generalized functions contains the tempered distributions $\mathcal{S}'(\mathbf{R}^n)$.

U. Ultradistributions

The localizing property of distributions is proved mainly from the existence of \dagger partitions of unity by functions in \mathcal{D} . Therefore if a test function space admits partitions of unity, the corresponding class of generalized functions

forms a sheaf. Roumieu [5] took the space $\mathcal{D}_{\{M_p\}}(\Omega)$ of \dagger ultradifferentiable functions of class $\{M_p\}$ with compact support. We can also use the space $\mathcal{D}_{(M_p)}(\Omega)$ of class (M_p) . The corresponding generalized functions are called **ultradistributions of class $\{M_p\}$** and of class (M_p) , respectively. Here M_p is a sequence of positive numbers satisfying the logarithmic convexity $M_p^2 \leq M_{p-1} M_{p+1}$ and the Denjoy-Carleman condition $\sum M_p / M_{p+1} < \infty$. $\mathcal{D}_{\{M_p\}}(\Omega)$ (resp. $\mathcal{D}_{(M_p)}(\Omega)$) is the space of all functions $\varphi \in \mathcal{D}(\Omega)$ such that there are constants k and C (resp. for any $k > 0$ there is a C) for which we have $|D^s \varphi(x)| \leq C k^{|s|} M_{|s|}$. In particular, the **Gevrey classes** $\{s\} = \{p^{1/s}\}$ and $(s) = (p^{1/s})$ defined for $s > 1$ are important, and they appear often in theory of differential equations. Almost all results for distributions have been extended to ultradistributions under appropriate conditions on M_p , which are satisfied by Gevrey sequences $p^{1/s}$ (\rightarrow [14] in particular). Closely connected are A. Beurling's test function spaces $\mathcal{D}_\omega(\Omega)$ (\rightarrow [13]), where ω is a function on \mathbf{R}^n continuous at the origin, satisfying $0 = \omega(0) \leq \omega(\xi + \eta) \leq \omega(\xi) + \omega(\eta)$ for all ξ and η in \mathbf{R}^n and such that $\int \omega(\xi)(1 + |\xi|)^{-n-1} d\xi < \infty$. $\mathcal{D}_\omega(\Omega)$ is defined to be the space of all $\varphi \in \mathcal{D}(\Omega)$ such that the Fourier transform $\hat{\varphi}$ satisfies $\int |\hat{\varphi}(\xi)| \exp(\lambda \omega(\xi)) d\xi < \infty$ for any $\lambda > 0$. If $\omega(\xi) = \log(1 + |\xi|)$, then $\mathcal{D}_\omega = \mathcal{D}$; if $\omega(\xi) = |\xi|^{1/s}$, then $\mathcal{D}_\omega = \mathcal{D}_{(s)}$. Let $\tilde{\omega}(\xi) = \omega(-\xi)$. The space $\mathcal{D}'_\omega(\Omega)$ of **Beurling's generalized distributions** is defined to be the strong dual of $\mathcal{D}_\omega(\Omega)$.

V. Hyperfunctions

In this and the following sections we mean by a cone $\Gamma \subset \mathbf{R}^n$ a convex open cone with vertex at 0. For two cones Γ, Δ , we write $\Delta \in \Gamma$ if $\Delta \cap \mathbf{S}^{n-1}$ is relatively compact in $\Gamma \cap \mathbf{S}^{n-1}$, where $\mathbf{S}^{n-1} = \{x \mid |x| = 1\}$. By a **wedge** we mean an open subset of \mathbf{C}^n of the form $\Omega + i\Gamma$, where $\Omega \subset \mathbf{R}^n$ is an open subset and $\Gamma \subset \mathbf{R}^n$ is a cone. Γ is called the opening of the wedge and Ω its edge. By an **infinitesimal wedge** (0-wedge for short) or a **tuboid** of opening Γ and edge Ω , we mean a complex open set U such that $U \subset \Omega + i\Gamma$ and that for any $\Delta \in \Gamma$, U contains the part of $\Omega + i\Delta$ which is contained in some complex neighborhood of the edge Ω . The symbol $\Omega + i\Gamma_0$ will represent any one of such open sets, and $\mathcal{O}(\Omega + i\Gamma_0)$ (the inductive limit of) the totality of functions holomorphic on some of them.

A **hyperfunction** $f(x)$ on an open set $\Omega \subset \mathbf{R}^n$ is an equivalence class of formal expressions of the form

$$f(x) = \sum_{j=1}^N F_j(x + i\Gamma_j 0), \quad (13)$$

where $F_j(z) \in \mathcal{O}(\Omega + i\Gamma_j 0)$. $\{F_j(z)\}$ is called a set

of **defining functions** of $f(x)$. Here the equivalence relation is given as

$$F_j(x + i\Gamma_j 0) + F_k(x + i\Gamma_k 0) = (F_j + F_k)(x + i\Gamma_j \cap \Gamma_k 0) \quad (14)$$

if $\Gamma_j \cap \Gamma_k \neq \emptyset$, that is, we can contract two terms into a single term as above and, conversely, can decompose, if possible, a term in the inverse way. These are considered to be modifications of the expression of the same hyperfunction. The totality of hyperfunctions on Ω is denoted by $\mathcal{B}(\Omega)$. It is a linear space by virtue of the linear structure naturally induced from that of holomorphic functions, combined with the above equivalence relation.

The symbol $F_j(x + i\Gamma_j 0)$, which represents by itself a hyperfunction, is called the **boundary value** of $F_j(z)$ to the real axis. It is merely formal and does not imply any topological limit, though there is some justification for the terminology as will be seen in Section Z.

In the case of one variable, we have only two kinds of wedges $\Omega + i\mathbf{R}^\pm$, hence a hyperfunction can be expressed by two terms:

$$F_+(x + i0) - F_-(x - i0). \quad (15)$$

Some examples of hyperfunctions of one variable are $\delta(x) = -(2\pi i)^{-1}((x + i0)^{-1} - (x - i0)^{-1})$; $\mathbf{1}(x) = -(2\pi i)^{-1}(\log(-x - i0) - \log(-x + i0))$; $\text{Pf } x^{-1} = ((x + i0)^{-1} + (x - i0)^{-1})/2$.

W. Localization of Hyperfunctions

If $\Omega' \subset \Omega$ is an open subset, the restriction mapping $\mathcal{B}(\Omega) \rightarrow \mathcal{B}(\Omega')$ is induced from that for holomorphic functions. With this structure the correspondence $\Omega \mapsto \mathcal{B}(\Omega)$ becomes a * presheaf. It is in fact a sheaf, because it can be expressed by the terminology of relative (or local) cohomology as follows: Let $H_\Omega^k(\mathbf{C}^n, \mathcal{F})$ denote the k th **relative cohomology group** of the pair $(\mathbf{C}^n, \mathbf{C}^n \setminus \Omega)$ (also called the k th **local cohomology group** with support in Ω) with coefficients in a sheaf \mathcal{F} on \mathbf{C}^n . (It is by definition the k th * derived functor of $\mathcal{F} \mapsto \Gamma_\Omega(\mathbf{C}^n, \mathcal{F}) =$ the totality of sections of \mathcal{F} defined on a neighborhood of Ω and with support in Ω and is calculated as the k th cohomology group of the * complex $\Gamma_\Omega(\mathbf{C}^n, \mathcal{L})$, where \mathcal{L} denotes any **flabby resolution** (i.e., * resolution by * flabby sheaves) of \mathcal{F} .) Let $\mathcal{H}_{\mathbf{R}^n}^k(\mathcal{F})$ denote the k th **derived sheaf** of \mathcal{F} to \mathbf{R}^n . (It is by definition the sheaf on \mathbf{R}^n associated with the presheaf $\Omega \mapsto H_\Omega^k(\mathbf{C}^n, \mathcal{F})$.) Then the cohomological definition of the sheaf of hyperfunctions is $\mathcal{B} = \mathcal{H}_{\mathbf{R}^n}^n(\mathcal{O})$ (the orientation being neglected). A fundamental theorem by Sato says that $\mathbf{R}^n \subset \mathbf{C}^n$ is **purely n -codimensional** with respect to \mathcal{O} (i.e., $\mathcal{H}_{\mathbf{R}^n}^k(\mathcal{O}) = 0$ for $k \neq n$), and moreover $H_\Omega^k(\mathbf{C}^n, \mathcal{O}) = 0$ for $k \neq n$ for any open set $\Omega \subset \mathbf{R}^n$.

Then by the general theory it can be shown that the remaining $H_\Omega^n(\mathbf{C}^n, \mathcal{O})$ agrees with the section module $\mathcal{B}(\Omega)$ of $\mathcal{H}_{\mathbf{R}^n}^n(\mathcal{O})$. Since $H^n(V, \mathcal{O}) = 0$ for any open set $V \subset \mathbf{C}^n$ (B. Malgrange), it follows further that the sheaf \mathcal{B} is flabby, i.e., its sections on any open set can always be extended to the whole space.

If U is a * Stein neighborhood of Ω , then $H_\Omega^n(\mathbf{C}^n, \mathcal{O})$ can be expressed using the covering cohomology as the quotient space

$$\mathcal{O}(U \# \Omega) \Big/ \sum_{j=1}^n \mathcal{O}(U \#_j \Omega), \quad (16)$$

where

$$\begin{aligned} U \# \Omega &= \{z \in U \mid \text{pr}_k(z) \notin \text{pr}_k(\Omega) \text{ for all } k\}, \\ U \#_j \Omega &= \{z \in U \mid \text{pr}_k(z) \notin \text{pr}_k(\Omega) \text{ for } k \neq j\}, \end{aligned} \quad (17)$$

and pr_k is the projection from \mathbf{C}^n to the k th coordinate. Then the isomorphism $H_\Omega^n(\mathbf{C}^n, \mathcal{O}) = \mathcal{B}(\Omega)$ is induced by the correspondence

$$\begin{aligned} \mathcal{O}(U \# \Omega) \ni F(z) &\mapsto [F(z)] \\ &= \sum_\sigma \text{sgn } \sigma F(x + i\Gamma_\sigma 0) \in \mathcal{B}(\Omega), \end{aligned} \quad (18)$$

where Γ_σ is the σ -orthant $\{y \in \mathbf{R}^n \mid \sigma_j y_j > 0, j = 1, \dots, n\}$ and $\text{sgn } \sigma = \sigma_1 \dots \sigma_n$. $F(z)$ is called a **defining function** of the corresponding hyperfunction. For one variable, any complex neighborhood $U \supset \Omega$ is Stein, and the above isomorphism reads $\mathcal{O}(U \setminus \Omega)/\mathcal{O}(U) = \mathcal{B}(\Omega)$, from which the naturality of the sign in (15) follows.

Thus the notion of support is also legitimate for hyperfunctions. The sheaf of hyperfunctions \mathcal{B} does not admit partitions of unity as for \mathcal{D}' . It is, however, flabby. Consequently, given a decomposition of a closed set into locally finite closed subsets $E = \bigcup_{\lambda \in \Lambda} E_\lambda$ and a hyperfunction f with support in E , we can always find hyperfunctions f_λ with support in E_λ such that $f = \sum_{\lambda \in \Lambda} f_\lambda$. For distributions this property holds only under some regularity assumption for the decomposition.

There are several practical criteria to determine whether or not a hyperfunction is zero in some open set Ω . These are called the **edge of the wedge theorem**. A hyperfunction $F(x + i\Gamma 0)$ with single expression is zero if and only if $F(z)$ itself is zero. $F_1(x + i\Gamma_1 0) = F_2(x + i\Gamma_2 0)$ if and only if they stick together to a function in $\mathcal{O}(\Omega + i(\Gamma_1 + \Gamma_2)0)$ (Epstein type). (Note that $\Gamma_1 + \Gamma_2$ is equal to the convex hull of $\Gamma_1 \cup \Gamma_2$, e.g., $\Gamma + (-\Gamma) = \mathbf{R}^n$ (Bogolyubov type).) $\sum_{j=1}^N F_j(x + i\Gamma_j 0) = 0$ if and only if there exist $G_{jk}(z) \in \mathcal{O}(\Omega + i(\Gamma_j + \Gamma_k)0)$, $j, k = 1, \dots, N$, such that $G_{jk}(z) = G_{kj}(z)$ and $F_j = \sum_{k=1}^N G_{jk}$, $j = 1, \dots, N$ (A. Martineau [19]). These are interpretations of cohomology in terms of coverings and have global variants concerning the envelope of holomorphy.

The real analytic functions $\varphi \in \mathcal{A}(\Omega)$ on Ω are naturally included in $\mathcal{B}(\Omega)$ via the ex-

Distributions and Hyperfunctions

pression $\varphi(x + i\Gamma 0)$ for any Γ or with $\Gamma = \mathbf{R}^n$. They form a subsheaf. Let f be a hyperfunction on Ω . The complement in Ω of the largest open subset $\Omega' \subset \Omega$, where $f(x)$ is real analytic, is called the **singular support** of $f(x)$ and is denoted by $\text{singsupp } f$. If Ω is bounded and $\text{singsupp } f \subset K$, then we can choose an expression of the form (13), where $F_j(z)$ can be continued analytically to $\Omega \setminus K$. If $\text{supp } f \subset K$, then these $F_j(z)$ satisfy $\sum F_j(x) = 0$ on $\Omega \setminus K$ in the usual sense.

X. Operations on Hyperfunctions

The **derivatives** of a hyperfunction $f(x)$ with the expression (13) are defined through the defining functions as $D^p f(x) = \sum (D_z^p F_j)(x + i\Gamma_j 0)$. The **product** by a real analytic function $\psi(x)$ is defined by $\psi(x)f(x) = \sum (\psi F_j)(x + i\Gamma_j 0)$. Combining these, we have the operation of a linear partial differential operator with real analytic coefficients $P(x, D)$ on hyperfunctions. It is a \dagger sheaf homomorphism.

Let $f \in \mathcal{B}(\Omega)$ and $D \subset \Omega$ be a compact set with piecewise smooth boundary. If $\text{singsupp } f \cap \partial D = \emptyset$, then by means of the special expression (13) mentioned at the end of Section W the **definite integral** is defined as

$$\int_D f(x) dx = \sum_{j=1}^N \int_{D_j} F_j(z) dz,$$

where D_j is a path deformed from D in such a way that $\partial D_j = \partial D$ and $F_j(z)$ is holomorphic on D_j . The result is independent of the choice of deformations or of the boundary value expression employed.

If $\text{supp } f \subset D$, the result is also independent of D ; hence it can be written as $\int_{\mathbf{R}^n} f(x) dx$. If $f(x, t)$ is a hyperfunction of the two groups of variables $(x, t) \in \Omega \times V$ such that $\text{singsupp } f \cap \partial D \times V = \emptyset$, then the integral $\int_D f(x, t) dx \in \mathcal{B}_t(V)$ is defined by the same method. It commutes with differentiation or integration with respect to t .

If $f(x) = \sum F_j(x + i\Gamma_j 0)$, $g(x) = \sum G_k(x + i\Delta_k 0)$, then we can define the **product** by $f(x)g(x) = \sum (F_j G_k)(x + i\Gamma_j \cap \Delta_k 0)$ under the assumption that $\Gamma_j \cap \Delta_k \neq \emptyset$ for every pair j, k . Especially, the product $f(x)g(t)$ is always legitimate for two hyperfunctions depending on different groups of variables. (It can also be interpreted as the tensor product.)

A real analytic coordinate transformation $x = \Phi(\tilde{x}) : \tilde{\Omega} \rightarrow \Omega$ extends naturally to a holomorphic coordinate transformation $z = \Phi(\tilde{z})$ on a complex neighborhood. A 0-wedge $\Omega + i\Gamma 0$ is transformed by Φ^{-1} to a twisted wedge containing, for each $x \in \Omega$ and $\Delta \in \Gamma$, a 0-wedge $\tilde{\Omega}_{\tilde{x}} + i(D\Phi^{-1})_x \Delta 0$ with some (real) neighborhood $\tilde{\Omega}_{\tilde{x}}$ of $\tilde{x} = \Phi^{-1}(x)$, where $(D\Phi^{-1})_x$ is the

derivative or the Jacobian matrix of the mapping Φ^{-1} at x . Thus the pullback $\Phi^* f(\tilde{x}) = f(\Phi(\tilde{x})) \in \mathcal{B}(\tilde{\Omega})$ of $f(x) \in \mathcal{B}(\Omega)$ by the transformation $x = \Phi(\tilde{x})$ can be defined by substituting $z = \Phi(\tilde{z})$ into the defining functions. It is consistent with the definition of the coordinate transformation for real analytic functions. Also, the law of the change of variables for the definite integral is the same as that for real analytic functions. This can be extended to the general operation of **substitution** as mentioned in Section Q for distributions or even to much more general operations (\rightarrow Section CC).

The **convolution** $(f * g)(x) = \int_{\mathbf{R}^n} f(x - t)g(t) dt$ is defined under the same assumption on support as in the case of distributions. It can be defined either literally as the composition of the above operations, or directly by way of the defining functions: For example, if g has compact support $\subset D$ and $\sum G_k(x + i\Delta_k 0)$ is an expression as mentioned at the end of Section W, then by choosing a suitable deformation D_k of D , we have

$$(f * g)(x) = \sum_{j,k} \left[\int_{D_k} F_j(z - \tau) G_k(\tau) d\tau \right]_{z \rightarrow x + i(\Gamma_j + \Delta_k) 0}.$$

Y. Hyperfunctions and Analytic Functionals

The totality $\mathcal{B}[K]$ of hyperfunctions with support in a compact set K becomes a nuclear Fréchet space. It is the dual of the nuclear (DF)-space $\mathcal{A}(K)$ of real analytic functions defined on a neighborhood of K . Thus $\mathcal{B}[K]$ is the space of analytic functionals with \dagger porter in the real compact set K . The duality is given by the definite integral $\langle f, \varphi \rangle = \int_{\mathbf{R}^n} f(x)\varphi(x) dx$ for $f \in \mathcal{B}[K]$ and $\varphi \in \mathcal{A}(K)$. A sequence $\{f_k(x)\}$ in $\mathcal{B}[K]$ converges if and only if it admits an expression (13) for a common fixed set of 0-wedges $\Omega + i\Gamma_j 0$, $j = 1, \dots, N$, such that the defining functions $F_{kj} \in \mathcal{O}(\Omega + i\Gamma_j 0)$ are also holomorphic on a fixed complex neighborhood U of $\Omega \setminus K$ and converges locally uniformly in $\Omega + i\Gamma_j 0 \cup U$, where Ω is a (real) neighborhood of K .

$\mathcal{B}[K]$ is isomorphic to $H_K^n(\mathbf{C}^n, \mathcal{O})$, and the above duality is a special case of the **Martineau-Harvey duality** $H_K^n(\mathbf{C}^n, \mathcal{O}) = \mathcal{O}(K)'$ for a Stein compact set $K \subset \mathbf{C}^n$. In the 1-dimensional case this is due to G. Köthe. If $K = [a_1, b_1] \times \dots \times [a_n, b_n]$, then $H_K^n(\mathbf{C}^n, \mathcal{O})$ can be represented, including the topology, by the quotient space

$$\mathcal{O}(U \# K) \Big/ \sum_{j=1}^n \mathcal{O}(U \#_j K),$$

where U is a Stein neighborhood of K and $U \# K$, $U \#_j K$ are defined in the same way as (17). If we choose $U = U_1 \times \dots \times U_n$, then by

way of a defining function $F(z) \in \mathcal{O}(U \setminus K)$ for $f \in \mathcal{B}[K]$ the inner product is given by the contour integral

$$\int_{\mathbb{R}^n} f(x) \varphi(x) dx = (-1)^n \oint_{\gamma_1} \dots \oint_{\gamma_n} F(z) \varphi(z) dz_1 \dots dz_n,$$

where $\gamma_j \subset U_j$ is a closed path surrounding $[a_j, b_j]$ once in the positive sense. Similar integral formulas are known for some special K of various types.

Starting from analytic functionals we can reconstruct the sheaf of hyperfunctions. For example, we can put $\mathcal{B}(\Omega)$ = the totality of locally finite sums of analytic functionals with support in Ω modulo the rearrangement of supports by decomposition (Martineau [17]). If Ω is bounded, we can also put $\mathcal{B}(\Omega) = \mathcal{B}[\bar{\Omega}] / \mathcal{B}[\partial\Omega]$ (Schapira [18]). The proof of localizability and/or flabbiness is based on the decomposability of support (which is the dual of the exact sequence $0 \rightarrow \mathcal{A}(K \cup L) \rightarrow \mathcal{A}(K) \oplus \mathcal{A}(L) \rightarrow \mathcal{A}(K \cap L) \rightarrow 0$) and the denseness of $\mathcal{B}[K] \subset \mathcal{B}[L]$ (which is the dual of the unique continuation property $0 \rightarrow \mathcal{A}(L) \rightarrow \mathcal{A}(K)$) for a pair $K \subset L$ with the same family of connected components. Note that in no way is the topology of $\mathcal{B}[K]$ localizable, or equivalently, $\mathcal{B}(\Omega)$ does not admit a reasonable topology.

Z. Embedding of Distributions

As the dual of the natural mapping $\mathcal{A}(K) \rightarrow \mathcal{E}(K)$ we have the topological embedding $\mathcal{E}'(K) \hookrightarrow \mathcal{A}'(K) = \mathcal{B}[K]$. This embedding conserves the support and hence gives rise to an embedding of sheaf $\mathcal{D}' \hookrightarrow \mathcal{B}$ (R. Harvey).

For a distribution T with compact support a set of its defining functions as a hyperfunction is given by $F_\sigma(z) = T_x(W(z-x, \Gamma_\sigma))$, σ being the multisignature, where $W(z, \Gamma_\sigma) = \int_{S^{n-1} \cap \Gamma_\sigma} W(z, \omega) d\omega$ and $W(z, \omega)$ is the component of a Radon decomposition of $\delta(x)$ (\rightarrow Section CC). If $\text{supp } T \subset K = [a_1, b_1] \times \dots \times [a_n, b_n]$, then as a hyperfunction it is represented by the following element of $\mathcal{O}(\mathbb{C}^n \setminus K)$:

$$F(z) = T_x \left(\frac{1}{(2\pi i)^n (x_1 - z_1) \dots (x_n - z_n)} \right). \tag{19}$$

It is in fact in $\mathcal{O}((\mathbb{P}^1)^n \setminus K)$ and vanishes at infinity, where $\mathbb{P}^1 = \mathbb{C}^1 \cup \{\infty\}$ is the Riemann sphere. These formulas are valid also for hyperfunctions, and they give defining functions of some canonical types. Especially, the one given by (19) is called the **standard defining function** and is characterized by the foregoing properties.

For the above-mentioned defining functions

for a distribution T , we have the following convergence in the sense of distributions on \mathbb{R}^n :

$$T_x = \sum_{\sigma} \text{sgn } \sigma \lim_{\epsilon \downarrow 0} F_{\sigma}(x + i\epsilon y_{\sigma}) \quad \text{for } y_{\sigma} \in \Gamma_{\sigma}.$$

More generally if the limit

$$\lim_{\epsilon \downarrow 0} F_j(x + i\epsilon y_j) \quad \text{for } y_j \in \Gamma_j \tag{20}$$

exists in $\mathcal{D}'(\Omega)$, then the distribution defined as this limit admits $\{F_j(z)\}$ as a set of defining functions when it is considered as a hyperfunction. However, for an arbitrary set of defining functions for a distribution, (20) does not necessarily converge in $\mathcal{D}'(\Omega)$. (This is because we can add terms with any bad behavior which cancel each other formally.) If a distribution admits the boundary value expression with only one term, or if the dimension is one, then the convergence of (20) in $\mathcal{D}'(\Omega)$ necessarily holds.

The convergence of (20) in $\mathcal{D}'(\Omega)$ is equivalent to the locally uniform estimate for the defining functions of the type $F_j(z) = O(|y|^{-M})$ for some $M > 0$. These assertions can be generalized to ultradistributions. For $\mathcal{D}'_{(M_p)}(\Omega)$ (resp. $\mathcal{D}'_{(M_p)}(\Omega)$) the last growth condition for the defining functions reads as follows: $F_j(z) = O(\exp M^*(L/|y|))$ for some L (resp. every $L > 0$), where $M^*(\rho) = \sup_p \log(\rho^p p! M_0 / M_p)$; especially for $M_p = p!^s$ we have $M^*(\rho) \sim \rho^{1/(s-1)}$.

AA. Hyperfunctions on a Real Analytic Manifold

Sticking the hyperfunctions on coordinate patches by the transformation law mentioned in Section X, we can define the sheaf of hyperfunctions on a real analytic manifold. More generally, for a real analytic vector bundle over a real analytic manifold, we can consider the sheaf of its hyperfunction local cross sections, which is also flabby. Thus, especially on a real analytic manifold M , we can obtain a concrete flabby resolution of the constant sheaf \mathbf{C}_M of length $\dim M$ by the sheaves of differential forms with hyperfunction coefficients: $0 \rightarrow \mathbf{C}_M \rightarrow \mathcal{B}_M^{(0)} \xrightarrow{d} \mathcal{B}_M^{(1)} \xrightarrow{d} \dots \rightarrow \mathcal{B}_M^{(\dim M)} \rightarrow 0$. With this sequence we can calculate the relative cohomology groups of open pairs with coefficients in \mathbb{C} by an analytic method. This is an extension of the de Rham theory for distributions [16].

If M is a compact manifold equipped with a nowhere vanishing real analytic density globally defined on M , then we have the topological duality $\mathcal{B}(M) = \mathcal{A}(M)'$. The inner product is given by the definite integral with respect to the density.

The Fourier series is an example of hyperfunctions on real analytic manifolds. The series $\sum c_p \exp(2\pi i p x)$ converges in $\mathcal{B}(\mathbf{T}^n)$ and defines a periodic hyperfunction if and only if c_p is of **infra-exponential growth**, i.e., $c_p = O(e^{\varepsilon|p|})$ for any $\varepsilon > 0$. \mathbf{T}^n has the global complex neighborhood $(\mathbf{P}^1)^n$, and $f(x) \in \mathcal{B}(\mathbf{T}^n)$ has the corresponding boundary value expression

$$F_\sigma(z) \in \mathcal{O}(\{|z_1|^{\sigma_1} < 1, \dots, |z_n|^{\sigma_n} < 1\}),$$

which represents the terms in the Fourier series such that $\sigma_j p_j > 0$.

BB. Fourier Hyperfunctions

In place of \mathcal{S} we take as the basic space the space \mathcal{P}_* of **exponentially decreasing real analytic functions** in the sense of M. Sato [15]: $f(x) \in \mathcal{P}_*$ if and only if there exist $\delta > 0$ and $\varepsilon > 0$ such that for any $\delta' < \delta$ and $\varepsilon' < \varepsilon$, $f(x)$ extends holomorphically to the neighborhood $\{\operatorname{Im} z| < \delta'\}$ of the real axis and has order $O(e^{-\varepsilon'|\operatorname{Re} z|})$ there. \mathcal{P}_* is endowed with the structure of nuclear (DF)-space via the inductive limit for $\delta > 0$, $\varepsilon > 0$. The classical Fourier transform \mathcal{F} acts isomorphically on \mathcal{P}_* . (In fact, δ and ε change their roles under \mathcal{F} .) The strong dual of \mathcal{P}_* is called the space of **Fourier hyperfunctions** and is denoted by \mathcal{Q} . It is a nuclear Fréchet space. It contains \mathcal{S}' as a dense subspace in view of the continuous and dense inclusion $\mathcal{P}_* \hookrightarrow \mathcal{S}$. It also contains classical locally integrable functions of **infra-exponential growth**, i.e., of order $e^{\varepsilon|x|}$ for any $\varepsilon > 0$. Thus by the duality we obtain a wider extension of Fourier transformation on \mathcal{Q} .

In the following a 0-wedge of the form $\mathbf{R}^n + i\Gamma 0$ will be called a 0-wedge of the form $\mathbf{D}^n + i\Gamma 0$ at the same time if it is a tubular domain (i.e., with fixed imaginary part $\Gamma 0$). Then an element $f(x) \in \mathcal{Q}$ can be expressed in the form (13), where each $F_j(z)$ is holomorphic in a 0-wedge $\mathbf{D}^n + i\Gamma_j 0$ and is of infra-exponential growth there locally uniformly in $\operatorname{Im} z$. The inner product of such $f(x)$ with $\varphi \in \mathcal{P}_*$ is given by the definite integral

$$\int f(x)\varphi(x)dx = \sum_{j=1}^N \int_{\operatorname{Im} z=y_j} F_j(z)\varphi(z)dz,$$

where the $y_j \in \Gamma_j 0$ are fixed. Given a cone Δ we define its **dual cone** by $\Delta^\circ = \{\eta \in \mathbf{R}^n | \langle \eta, y \rangle \geq 0 \text{ for all } y \in \Delta\}$. If $F_j(z)$ are all of exponential decrease in $\operatorname{Re} z$ locally uniformly with respect to $\operatorname{Im} z \in \Gamma_j 0$ and $\operatorname{Re} z/|\operatorname{Re} z| \notin \Delta^\circ$, then the definite integral

$$\begin{aligned} G(\zeta) &= (\sqrt{2\pi})^{-n} \int e^{-ix\zeta} f(x) dx \\ &= \sum_{j=1}^N (\sqrt{2\pi})^{-n} \int_{\operatorname{Im} z=y_j} e^{-iz\zeta} F_j(z) dz \end{aligned} \quad (21)$$

converges locally uniformly for ζ in some 0-wedge $\mathbf{D}^n - i\Delta 0$, and defines there a holomorphic function of infra-exponential growth. Thus we obtain a Fourier hyperfunction $G(\zeta - i\Delta 0)$ that agrees with $\mathcal{F}f$ calculated by the duality. For a general $f(x) \in \mathcal{Q}$, the Fourier transform in the manner of Sato is calculated as follows: First we decompose $f(x)$ into the sum $\sum f_k(x)$ for which the defining functions of $f_k(x)$ decrease exponentially outside Δ_k° . Then we calculate $G_k(\zeta)$ by (21) and put $\mathcal{F}f = \sum G_k(\zeta - i\Delta_k 0)$. An example of such decomposition is given by multiplication by $\chi_\sigma(x) = \prod_{j=1}^n 1/(1 + \exp \sigma_j x_j)$, which decreases exponentially outside $\Gamma_\sigma = \{\sigma_j x_j \geq 0\}$.

The relation between \mathcal{Q} and \mathcal{B} is more complicated than the relation $\mathcal{S}' \subset \mathcal{D}'(\mathbf{R}^n)$. The growth condition for \mathcal{Q} is interpreted as a condition concerning germs at infinity. Thus \mathcal{Q} can be considered to be a sheaf on the directional compactification $\mathbf{D}^n = \mathbf{R}^n \cup \mathbf{S}_\infty^{n-1}$ such that $\mathcal{Q}|_{\mathbf{R}^n} = \mathcal{B}$. Just as the sheaf \mathcal{B} is obtained from \mathcal{O} , the sheaf \mathcal{Q} is obtained as the n th derived sheaf $\mathcal{H}_{\mathbf{D}^n}^n(\tilde{\mathcal{O}})$ from the sheaf $\tilde{\mathcal{O}}$ on $\mathbf{D}^n + i\mathbf{R}^n$ consisting of germs of holomorphic functions of infra-exponential growth with respect to $\operatorname{Re} z$. We have $H_\Omega^k(\mathbf{D}^n + i\mathbf{R}^n, \tilde{\mathcal{O}}) = 0$ for $k \neq n$ for any open set $\Omega \subset \mathbf{D}^n$. Especially, \mathcal{Q} is flabby, and the decomposition of support is available to calculate the Fourier transform. The symbol \mathcal{Q} employed at the beginning to express the global Fourier hyperfunctions corresponds to $\mathcal{Q}(\mathbf{D}^n)$, and $\mathcal{Q}(\mathbf{R}^n) = \mathcal{B}(\mathbf{R}^n)$ by definition. The canonical restriction mapping $\mathcal{Q}(\mathbf{D}^n) \rightarrow \mathcal{B}(\mathbf{R}^n)$ is surjective but not injective.

As for tempered distributions \mathcal{S}' , we can introduce various subclasses of Fourier hyperfunctions, e.g., **exponentially decreasing Fourier hyperfunctions** $\bigcup_{\delta>0} \exp(-\delta\sqrt{x^2+1})\mathcal{Q}$, real analytic functions of infra-exponential growth $\tilde{\mathcal{O}}(\mathbf{D}^n)$, etc. We can also consider operations such as convolution and multiplication between adequate pairs, and apply differential operators with suitable coefficients. Concerning these we can avail ourselves of the same formulas as given in Section O.

A hyperfunction with compact support is naturally considered as a Fourier hyperfunction, and its Fourier transform agrees with the inner product $\langle f(x), (\sqrt{2\pi})^{-n} \exp(-ix\xi) \rangle$, which gives an entire function of exponential type.

Paley-Wiener theorem. An entire function $f(\zeta)$ is the Fourier transform of a hyperfunction with support in a compact convex set K in \mathbf{R}^n if and only if it satisfies condition (i) of Section O.

The theory of Fourier hyperfunctions described above is mainly due to Sato and T. Kawai [20]. They are not the largest class of generalized functions stable under the Fourier

transformation. Following a suggestion of Sato, S. Nagamachi and N. Mugibayashi, Y. Saburi, and Y. Ito have extended them to the **modified Fourier hyperfunctions** in which the radial compactification \mathbf{D}^{2n} of \mathbf{C}^n is employed instead of the horizontal compactification $\mathbf{D}^n + i\mathbf{R}^n$ in the above theory. If we discard the localizing property, they can be extended further to the **Fourier ultrahyperfunctions** or **ultradistributions** of J. Sebastião e Silva (1958), M. Hasumi (1961), and M. Morimoto (1973).

CC. Micro-Analyticity of Hyperfunctions

The boundary value expression (13) for a hyperfunction $f(x)$ can be interpreted conversely as the description of the state of analytic continuation of $f(x)$ to the complex domain. Thus we say that f is **micro-analytic** at (x_0, ξ_0) if on a neighborhood of x_0 it admits the analytic continuation into the half-space $\langle \text{Im } z, \xi_0 \rangle < 0$ in the sense that it admits an expression (13) satisfying $\Gamma_j \cap \{ \langle \text{Im } z, \xi_0 \rangle < 0 \} \neq \emptyset$ for every j . The set of points $(x_0, \xi_0) \in \Omega \times S^{n-1}$, where $f \in \mathcal{B}(\Omega)$ is not micro-analytic, is called the **singularity spectrum** or **singular spectrum** of f and is denoted by S.S. f . We have by definition $\text{S.S.}F(x + i\Gamma 0) \subset \Omega \times (\Gamma^0 \cap S^{n-1})$. Micro-analyticity can be characterized by the Fourier transformation as follows: f is micro-analytic at (x_0, ξ_0) if and only if there exists a Fourier hyperfunction $g(x)$ such that the Fourier transform $\hat{g}(\xi)$ decreases exponentially on a conical neighborhood of ξ_0 and that $f - g$ is real analytic on a neighborhood of x_0 . A hyperfunction $f(x)$ is real analytic in a neighborhood of x_0 if and only if it is micro-analytic at (x_0, ξ) for any $\xi \in S^{n-1}$.

With this notion we can clarify the operations on hyperfunctions. (In the following S^{n-1} is identified with $(\mathbf{R}^n \setminus \{0\})/\mathbf{R}^+$ and $+$ denotes the sum in the latter.) We have

$$\text{S.S.}(fg) \subset \{ (x, \lambda\xi + (1-\lambda)\eta) \mid (x, \xi) \in \text{S.S.}f, (x, \eta) \in \text{S.S.}g, 0 \leq \lambda \leq 1 \} \cup \text{S.S.}f \cup \text{S.S.}g,$$

$$\text{S.S.}f(\Phi(\tilde{x})) \subset \{ (\tilde{x}, {}^tD\Phi\xi) \mid (\Phi(\tilde{x}), \xi) \in \text{S.S.}f \},$$

and these operations are legitimate if and only if 0 does not appear in the direction component of the right-hand side. We also have

$$\text{S.S.} \int f(x, t) dx \subset \{ (x, \xi) \mid (x, t, \xi, 0) \in \text{S.S.}f \text{ for a } t \},$$

$$\text{S.S.}(f * g) \subset \{ (x + y, \xi) \mid (x, \xi) \in \text{S.S.}f, (y, \xi) \in \text{S.S.}g \}$$

under suitable conditions for support.

The classical **decomposition formula of Radon** (or the **plane wave decomposition** of δ)

$$\delta(x) = \frac{(n-1)!}{(-2\pi i)^n} \int_{S^{n-1}} \frac{d\omega}{(x\omega + i0)^n} \tag{22}$$

can be interpreted as the decomposition of δ into hyperfunctions with S.S. in the single direction ω . By the convolution, this formula supplies similar decomposition for a general hyperfunction (called the singular spectrum decomposition). (22) can be generalized to

$$\delta(x) = \frac{(n-1)!}{(-2\pi i)^n} \int_{S^{n-1}} \frac{\det(\text{grad}_\omega \Psi(x, \omega))}{(\Phi(x, \omega) + i0)^n} d\omega, \tag{23}$$

where the twisted phase $\Phi(x, \omega)$ satisfies (i) $\Phi(x, \omega)$ is a real analytic function of positive type (i.e., $\text{Re } \Phi(x, \omega) = 0$ implies $\text{Im } \Phi(x, \omega) \geq 0$) and (ii) $\Phi(x, \omega)$ is homogeneous of order 1 in ω and $\Phi(0, \omega) = 0$, $\text{grad}_x \Phi(0, \omega) = \omega$; and the vector $\Psi(x, \omega)$ is such that $\langle \Psi(x, \omega), x \rangle = \Phi(x, \omega)$. If $\Phi(x, \omega)$ further satisfies $\Phi(x, \omega) \neq 0$ for $x \neq 0$, then the component becomes a hyperfunction (even a distribution) of x whose S.S. is precisely one point $(0, \omega)$, and this fact is useful in theoretical applications [7, 21] (\rightarrow 274 Microlocal Analysis).

DD. Structure Theorems of Hyperfunctions

A hyperfunction whose support is concentrated at the origin is expressed as the infinite derivative $J(D)\delta(x) = \sum a_p D^p \delta(x)$ of the Dirac measure, with the coefficients satisfying $\lim(|a_p|p!)^{1/p} = 0$. Such an operator $J(D)$ is called a **local operator** with constant coefficients and acts on \mathcal{B} as a sheaf homomorphism. Its total symbol $J(\zeta)$ is an entire function of infra-exponential growth or of type $[1, 0]$. By way of such an operator, a general (Fourier) hyperfunction can be written in the form $J(D)g(x)$ with a continuous function on \mathbf{R}^n (of infra-exponential growth).

If $0 \in \text{supp } f \subset \{ \langle v, x \rangle \geq 0 \}$, then $\text{S.S.}f \ni (0, \pm v)$ (**Holmgren type theorem** of Kashiwara and Kawai), and furthermore, the direction component of S.S. f at 0 has the form of $\{ \pm v \} \cup \rho^{-1}(E)$ with some $E \subset S^{n-2}$, where $\rho: S^{n-1} \setminus \{ \pm v \} \rightarrow S^{n-2}$ is the projection to the equator (the **watermelon-slicing theorem** of Morimoto, Kashiwara, and K. Kataoka). E is called the reduced S.S. of f at 0. These theorems have many applications in the theory of linear partial differential equations and also in physics.

A hyperfunction $f(x)$ with support in the hyperplane $x_n = 0$ has several further remarkable properties in x_n . It admits a formal expansion of the form $\sum_{k=0}^\infty f_k(x') \delta^{(k)}(x_n)$, where $x' = (x_1, \dots, x_{n-1})$ and $f_k(x') = \int_{-\infty}^\infty f(x) x_n^k / k! dx_n$. The sum converges in the sense of the topology if f has compact support. It reduces locally to a finite sum if f is a distribution, and the k th term represents the k -ple layer distribution of mass, electric charge, etc. For a general f the coefficients $\{ f_k(x') \}$ do not neces-

sarily determine f , though they are determined by f .

EE. Complex Powers of Polynomials

Among examples of special generalized functions the most important are those of the form f_+^λ , where $f_+ = \max\{f(x), 0\}$ (or more generally it can be replaced by zero on some connected components of $\{f(x) > 0\}$), and $\lambda \in \mathbb{C}$ denotes a holomorphic parameter. (The discussion is the same for $f_- = \max\{-f(x), 0\}$.) The simplest example, x_+^λ , is defined as the analytic continuation of the locally integrable function x_+^λ for $\operatorname{Re} \lambda > -1$ by repeated use of the formula $x_+^\lambda = (\lambda + 1)^{-1} D_x x_+^{\lambda+1}$, and becomes meromorphic in λ with simple poles at $\lambda = -1, -2, \dots$. As a hyperfunction we have $x_+^\lambda = \{(-x + i0)^\lambda - (-x - i0)^\lambda\}/2i \sin \pi \lambda$. At a negative integer $\lambda = -n$, x_+^λ has residue $(-1)^{n-1} \delta^{(n-1)}(x)/(n-1)!$ and finite part $[-(2\pi i)^{-1} z^{-n} \log(-z)]$. The latter is often denoted by x_+^{-n} . In general, for a germ of a real-valued real analytic function $f(x)$ we can find a differential operator $P(\lambda, x, D_x)$ with polynomial coefficients in λ and a monic polynomial $b(\lambda)$ of minimum degree such that

$$P(\lambda, x, D_x) f_+^{\lambda+1} = b(\lambda) f_+^\lambda \quad (24)$$

(Sato, I. N. Bernshtein, Kashiwara, J.-E. Björk [22]). This formula gives the analytic continuation of f_+^λ just as for x_+^λ . The polynomial $b(\lambda)$ is called the ***b*-function** or the **Sato-Bernshtein polynomial** and contains valuable information regarding the singularity of f . It has only negative rational roots (Kashiwara).

We have $f \cdot f_\pm^\lambda = f_\pm^{\lambda+1}$; hence $f_+^{-1} - f_-^{-1}$ (suitably interpreted as above) gives a solution of the division problem $u \cdot f = 1$. Thus if f is a polynomial, its inverse Fourier transform gives a tempered fundamental solution of $f(-iD)$. Furthermore, when f is the relative invariant of a * prehomogeneous vector space, we can calculate $b(\lambda)$ explicitly by way of the holonomy diagram. Also, the Fourier transform of f_+^λ can be calculated explicitly by way of the real holonomy diagram as a linear combination of the corresponding objects for the dual prehomogeneous vector space with coefficients similar to the * Maslov index. The simplest example is

$$\begin{aligned} \mathcal{F} x_\pm^\lambda &= (\sqrt{2\pi})^{-1} (\mp i) e^{-\pi i \lambda / 2} \Gamma(\lambda + 1) (\xi \mp i 0)^{-\lambda-1} \\ &= (\sqrt{2\pi})^{-1} (\mp i) \Gamma(\lambda + 1) \{e^{\mp \pi i \lambda / 2} \xi_-^{-\lambda-1} \\ &\quad - e^{\pm \pi i \lambda / 2} \xi_+^{-\lambda-1}\}. \end{aligned}$$

Among practical examples are the following classical formulas: Let $P(x)$ be a nondegenerate real quadratic form and $Q(\xi)$ its * dual form, and let q denote the number of negative

eigenvalues of $P(x)$. Then

$$\begin{aligned} \mathcal{F} P_+^\lambda &= -(\sqrt{2\pi})^{-n} 2^{2\lambda+n} \pi^{n/2-1} \Gamma(\lambda+1) \Gamma(\lambda+n/2) \\ &\quad \times |\det P|^{-1/2} \{\sin \pi(\lambda+q/2) Q_+^{-\lambda-n/2} \\ &\quad - \sin \pi((n-q)/2) Q_-^{-\lambda-n/2}\}. \end{aligned}$$

Here the arguments in the Γ -factor $(\lambda+1)(\lambda+n/2)$ give the b -function of $P(x)$. If $q = n-1$, we further have, letting $P_\pm^\lambda = P_\pm^\lambda \mathbf{1}(\pm \langle x, v \rangle)$ for an eigenvector v corresponding to the unique positive eigenvalue,

$$\begin{aligned} \mathcal{F} P_\pm^\lambda &= (2\pi)^{-n/2} 2^{2\lambda+n-1} \pi^{n/2-1} \Gamma(\lambda+1) \\ &\quad \times \Gamma(\lambda+n/2) |\det P|^{-1/2} \{e^{\mp \pi i(\lambda+n/2)} Q_\pm^{-\lambda-n/2} \\ &\quad + e^{\pm \pi i(\lambda+n/2)} Q_\mp^{-\lambda-n/2} + Q_-^{-\lambda-n/2}\}. \end{aligned}$$

From these formulas (taking the finite part if necessary) we obtain the fundamental solution of the wave equation, the Laplacian, and their iterations. These are exactly the distributions introduced by Hadamard, M. Riesz, and others, as mentioned in Section A (\rightarrow Appendix A, Table 15.V).

References

- [1] L. Schwartz, *Théorie des distributions*, Hermann, revised edition, 1966.
- [2] L. Hörmander, *Linear partial differential operators*, Springer, 1963.
- [3] G. de Rham, *Variétés différentiables*, Hermann, 1955.
- [4] I. M. Gel'fand and G. E. Shilov, *Generalized functions*. I, Properties and operations; II, Spaces of fundamental and generalized functions; III, Theory of differential equations; IV (I. M. Gel'fand and N. Ya. Vilenkin), Applications of harmonic analysis; V (I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin), Integral geometry and representation theory, Academic Press, 1964, 1968, 1967, 1964, 1966; VI (I. M. Gel'fand, M. I. Graev, and I. I. Pyatetskii-Shapiro), Representation theory and automorphic functions, Saunders, 1969. (Originals in Russian.)
- [5] C. Roumieu, Sur quelques extensions de la notion de distribution, *Ann. Sci. Ecole Norm. Sup. Paris*, 77 (1960), 41–121.
- [6] M. Sato, Theory of hyperfunctions I, II, *J. Fac. Sci. Univ. Tokyo*, sec. I, 8 (1959), 139–193, 387–437.
- [7] M. Sato, T. Kawai, and M. Kashiwara, Microfunctions and pseudo-differential equations, in [16, pp. 265–529].
- [8] L. Schwartz, Espaces de fonctions différentiables à valeurs vectorielles, *J. Analyse Math.*, 4 (1954–1955), 88–148.
- [9] A. Grothendieck, Produits tensoriels topologiques et espaces nucléaires, *Mém. Amer. Math. Soc.*, 16 (1955).

- [10] L. Schwartz, *Théorie des distributions à valeurs vectorielles*, Ann. Inst. Fourier, 7 (1957), 1–141; 8 (1958), 1–209.
- [11] F. Trèves, *Topological vector spaces, distributions and kernels*, Academic Press, 1967.
- [12] L. Ehrenpreis, *Fourier analysis in several complex variables*, Wiley-Interscience, 1970.
- [13] G. Björck, *Linear partial differential operators and generalized distributions*, Ark. Mat., 6 (1966), 351–407.
- [14] H. Komatsu, *Ultradistributions. I, Structure theorems and a characterization; II, The kernel theorem and ultradistributions with support in a submanifold; III, Vector-valued ultradistributions and the theory of kernels*, J. Fac. Sci. Univ. Tokyo, sec. IA, 20 (1973), 25–105; 24 (1977), 607–628; 29 (1982), 653–718.
- [15] M. Sato, *Theory of hyperfunctions* (in Japanese), Sugaku, 10 (1958), 1–27.
- [16] H. Komatsu (ed.), *Hyperfunctions and pseudo-differential equations*, Lecture notes in math. 287, Springer, 1973.
- [17] A. Martineau, *Les hyperfonctions de M. Sato*, Sémin. Bourbaki, 13 (1960–1961), no. 214.
- [18] P. Schapira, *Théorie des hyperfonctions*, Lecture notes in math. 126, Springer, 1970.
- [19] A. Martineau, *Le “edge of the wedge theorem” en théorie des hyperfonctions de Sato*, Proc. Intern. Conf. Functional Analysis and Related topics, Tokyo, 1969, Univ. Tokyo Press, 1970, 95–106.
- [20] T. Kawai, *On the theory of Fourier hyperfunctions and its applications to partial differential equations with constant coefficients*, J. Fac. Sci. Univ. Tokyo, sec. IA, 17 (1970), 467–517.
- [21] K. Kataoka, *On the theory of Radon transformations of hyperfunctions*, J. Fac. Sci. Univ. Tokyo, sec. IA, 28 (1981), 331–413.
- [22] J.-E. Björk, *Rings of differential operators*, North-Holland, 1979.

126 (IX.22) Dynamical Systems

A. History

The theory of dynamical systems began with the investigation of the motion of planets in ancient astronomy. Qualitative investigation of mechanics in antiquity and the Middle Ages culminated in the work of J. Kepler and G. Galilei in 17th century. At the end of that century, I. Newton founded his celebrated Newtonian mechanics, by means of which

Kepler's law on the motion of planets and Galileo's observations of movement can be explained theoretically. Following this, L. Euler, J. L. Lagrange, P. S. Laplace, W. R. Hamilton, C. G. J. Jacobi, and others developed the theory using analytical methods, and founded analytical dynamics. From the end of the 18th century through the 19th century, the [†]three-body problem attracted the attention of many mathematicians. At the end of the 19th century, H. Bruns and H. Poincaré found that general solutions of the three-body problem could not be obtained by [†]quadrature, and this gave rise to a crisis of analytical dynamics. But this was resolved by Poincaré himself. He pointed out the importance of the qualitative theory based on topological methods, and obtained many fundamental results. A. M. Lyapunov with his theory of stability and G. D. Birkhoff with his many important concepts of topological dynamics established foundations of the new qualitative theory.

In 1937 A. A. Andronov and L. S. Pontryagin introduced the concept of structural stability, which attracted the attention of S. Lefschetz. Lefschetz's school investigated structural stability and [†]nonlinear oscillations, and obtained many important results (H. F. de Baggis, L. Markus, M. M. Peixoto, and others). In about 1960, S. Smale initiated study of differentiable dynamical systems under the influence of Lefschetz's school. Smale and his school founded a new theory of differentiable dynamical systems using [†]differential topology. D. V. Anosov generalized the work of E. Hopf and G. A. Hedlund on [†]geodesic flows of closed surfaces of [†]constant negative curvature and established the concept of Anosov systems, which played an important role in Smale's theory. The work of Hopf, Hedlund, and Anosov is closely related to [†]ergodic theory. Ya. G. Sinai and R. Bowen obtained important results in ergodic theory. The concept of structural stability and its generalization are essential in the [†]catastrophe theory of R. Thom (→ 51 Catastrophe Theory); the theory of bifurcation of dynamical systems is another essential part of catastrophe theory. D. Ruelle and F. Takens proposed a new mathematical mechanism for the generation of turbulence using Smale's theory and Hopf bifurcation. The new theory of dynamical systems developed by Smale and others is now applied to the mathematical explanation of chaotic phenomena in many branches of science. Finally, we mention that in the 1960s A. N. Kolmogorov, V. I. Arnold, and J. Moser obtained remarkable results on the existence of quasi-periodic solutions for the *n*-body problem, which turned out to solve the long-standing problem of the stability of the solar system.

B. Definitions of Dynamical Systems

In the study of the evolution of physical, biological, and other systems, we construct mathematical models of the systems. Usually, the state of a given system is completely described by a collection of continuous parameters, which may be related in some cases. Thus the space X of all possible states of the system can be regarded as a Euclidean space or a subset of a Euclidean space defined by some equations. In general, we assume that the space X of all possible states of the system forms a [†]topological space, and we call it a **state space** or a **phase space**. Second, we assume that the law of evolution of states in time is given, by which we can tell the state x_1 at any time t_1 if we know the state x_0 at time t_0 . Assigning x_1 to x_0 , we have a mapping $\pi(t_1, t_0): X \rightarrow X$ for any times t_0 and t_1 , which satisfies the following conditions: (i) $\pi(t_2, t_1) \circ \pi(t_1, t_0) = \pi(t_2, t_0)$; (ii) $\pi(t_0, t_0) = 1_X$, the identity mapping of X . Finally, we assume that the mapping $\pi(t_1, t_0)$ depends only on $t = t_1 - t_0$. Writing $\pi_t = \pi(t_1, t_0)$ if $t = t_1 - t_0$, we have the following conditions from (i) and (ii) above: (i') $\pi_s \circ \pi_t = \pi_{s+t}$, $s, t \in \mathbf{R}$; (ii') $\pi_0 = 1_X$.

In general, the theory of topological dynamics can be regarded as the study of topological transformation groups (\rightarrow 431 Transformation Groups) originating in the topological investigations of problems arising from classical mechanics. Here, we restrict our attention to some important special cases.

(1) Let X be a topological space and \mathbf{R} the additive topological group of real numbers. Let $\varphi: X \times \mathbf{R} \rightarrow X$ be a continuous mapping. For each $t \in \mathbf{R}$, we define a mapping $\varphi_t: X \rightarrow X$ by $\varphi_t(x) = \varphi(x, t)$, $x \in X$. If the family of mappings $\{\varphi_t\}_{t \in \mathbf{R}}$ satisfies the following conditions, we say that (X, φ) is a **(continuous) \mathbf{R} -action**, a **(continuous) flow**, or a **(continuous) dynamical system** on X , and that X is the **phase space**: (i) $\varphi_s \circ \varphi_t = \varphi_{s+t}$ for all $s, t \in \mathbf{R}$; (ii) $\varphi_0 = 1_X$.

Let (X, φ) be a flow. Then $\varphi_t: X \rightarrow X$ is a [†]homeomorphism with $(\varphi_t)^{-1} = \varphi_{-t}$ for each $t \in \mathbf{R}$. For each $x \in X$, define a mapping $\varphi^x: \mathbf{R} \rightarrow X$ by $\varphi^x(t) = \varphi(x, t)$, $t \in \mathbf{R}$. The mapping φ^x is called a **motion** through x , and its image $C(x) = \{\varphi^x(t) | t \in \mathbf{R}\}$ is called the **orbit** or the **trajectory** through x . An orbit is nonempty, and two orbits are either identical or mutually disjoint. The family of orbits fills up the phase space X and is called the **phase portrait**.

Let \mathbf{R}_+ be the additive [†]semigroup of all nonnegative real numbers. If we replace \mathbf{R} by \mathbf{R}_+ in the definition of a (continuous) flow, we obtain a definition of a **(continuous) semiflow**. For a semiflow (X, φ) , the mapping $\varphi_t: X \rightarrow X$, $t \in \mathbf{R}_+$ is in general not a homeomorphism but a continuous mapping.

Let (X, φ) and (Y, ψ) be flows. A homeomorphism $h: X \rightarrow Y$ is called a **topological equivalence** from (X, φ) to (Y, ψ) if it maps each orbit of φ onto an orbit of ψ preserving orientations of orbits (i.e., there exists an increasing homeomorphism $\alpha_x: \mathbf{R} \rightarrow \mathbf{R}$ for each $x \in X$ such that $h\varphi(x, t) = \psi(h(x), \alpha_x(t))$ for all $t \in \mathbf{R}$). Two flows are **topologically equivalent** if there exists a topological equivalence from one to the other. If two flows are topologically equivalent, their phase portraits have the same topological structure. Two flows (X, φ) and (Y, ψ) are **flow equivalent** if there exist a $c > 0$ and a homeomorphism $h: X \rightarrow Y$ such that $h\varphi(x, t) = \psi(h(x), ct)$ for all $t \in \mathbf{R}$. Such an h is a topological equivalence from (X, φ) to (Y, ψ) .

(2) Let X be a topological space and \mathbf{Z} the additive group of integers. If we replace \mathbf{R} by \mathbf{Z} in the definition of a continuous flow, we obtain a definition of a **(continuous) \mathbf{Z} -action**, a **discrete flow**, or a **discrete dynamical system** on X . If (X, φ) is a discrete flow, then $f = \varphi_1: X \rightarrow X$ is a homeomorphism and $\varphi_n = f^n$ for all $n \in \mathbf{Z}$. Conversely, for a given homeomorphism $f: X \rightarrow X$, define a mapping $\varphi: X \times \mathbf{Z} \rightarrow X$ by $\varphi(x, n) = f^n(x)$, $x \in X$ and $n \in \mathbf{Z}$. Then (X, φ) is a discrete flow such that $\varphi_n = f^n$ for $n \in \mathbf{Z}$. So we identify a homeomorphism with a discrete flow. Thus the orbit of a homeomorphism $f: X \rightarrow X$ through x is $C(x) = \{f^n(x) | n \in \mathbf{Z}\}$.

Let \mathbf{Z}_+ be the additive semigroup of all nonnegative integers. If we replace \mathbf{R} by \mathbf{Z}_+ in the definition of a continuous flow, we obtain a definition of a **discrete semiflow** on X . For a discrete semiflow (X, φ) , the mapping $\varphi_n: X \rightarrow X$, $n \in \mathbf{Z}_+$ is in general not a homeomorphism but a continuous mapping. We can identify a continuous mapping $f: X \rightarrow X$ with a discrete semiflow (X, φ) in a natural way as above.

Let $f: X \rightarrow X$ and $g: Y \rightarrow Y$ be two homeomorphisms (continuous mappings). A homeomorphism $h: X \rightarrow Y$ such that $h \circ f = g \circ h$ is called a **topological conjugacy** from f to g . And f and g are called **topologically conjugate** if there exists a topological conjugacy from f to g . Topologically conjugate homeomorphisms have the same phase portrait in a topological sense.

(3) Let M be a [†]differentiable manifold of class C^r ($1 \leq r \leq \infty$ or $r = \omega$). A continuous flow (M, φ) is a **flow of class C^r** , a **C^r -flow**, a **differentiable dynamical system of class C^r** , or a **one-parameter group of transformations of class C^r** , if $\varphi: M \times \mathbf{R} \rightarrow M$ is of class C^r . A **semiflow of class C^r** is defined similarly.

Let (M, φ) and (N, ψ) be C^r -flows. A topological equivalence $h: M \rightarrow N$ from (M, φ) to (N, ψ) is called a **C^r -equivalence** if it is a [†] C^r -diffeomorphism. Two flows are **C^r -equivalent** if there is a C^r -equivalence from one to the other. Classification of C^r -flows by C^r -equivalence is

difficult and sometimes too unwieldy to work with. On the other hand, there are many problems which can be solved by the knowledge of the topological structure of the phase portrait of C^r -flows.

(4) Let (M, π) be a discrete flow on a differentiable manifold M of class C^r . If $\pi: M \times \mathbb{Z} \rightarrow M$ is of class C^r , then (M, π) is called a **Z-action of class C^r** , a **discrete flow of class C^r** , a **discrete C^r -flow**, or a **discrete dynamical system of class C^r** . If (M, π) is a discrete C^r -flow, then $f = \pi_1: M \rightarrow M$ is a C^r -diffeomorphism. Conversely, a C^r -diffeomorphism $f: M \rightarrow M$ defines a discrete C^r -flow in a natural way on M , and we identify a C^r -diffeomorphism $f: M \rightarrow M$ with the discrete C^r -flow on M defined by f . A **discrete semiflow of class C^r** is defined similarly, and a C^r -mapping $f: M \rightarrow M$ is identified with a discrete semiflow of class C^r on M defined by f .

Let $f: M \rightarrow M$ and $g: N \rightarrow N$ be C^r -diffeomorphisms (C^r -mappings) of differentiable manifolds M and N of class C^r . A topological conjugacy from f to g is called a **C^r -conjugacy** if it is a C^r -diffeomorphism. f and g are **C^r -conjugate** if there is a C^r -conjugacy from f to g .

C. Examples and Remarks

(1) Let D be an open set of \mathbb{R}^n and $f: D \rightarrow \mathbb{R}^n$ a continuous mapping. Consider the autonomous system of ordinary differential equations

$$dx/dt = f(x), \quad x \in D. \quad (1)$$

We assume that for each $x \in D$ there exists a unique nonextendable solution $\varphi(x, t)$ with the initial condition $\varphi(x, 0) = x$ defined on a maximal interval (a_x, b_x) , $-\infty \leq a_x < 0 < b_x \leq \infty$ (\rightarrow 316 Ordinary Differential Equations (Initial Value Problems)). The set $\{\varphi(x, t) | a_x < t < b_x\}$ is called the **trajectory** through x . By the uniqueness assumption, we have $\varphi(\varphi(x, t), s) = \varphi(x, t + s)$ whenever both sides of the equality are defined. When $(a_x, b_x) = \mathbb{R}$ for all $x \in D$, then equation (1) is called **complete**. If (1) is complete, the mapping $\varphi: D \times \mathbb{R} \rightarrow D$ defined by the solutions $\varphi(x, t)$ determines a continuous flow (D, φ) . Furthermore, if f is of class C^r , then (D, φ) is of class C^r . If (1) is not complete, then there exists a continuous positive scalar function $\alpha: D \rightarrow \mathbb{R}$ such that

$$dx/dt = \alpha(x)f(x), \quad x \in D, \quad (2)$$

is complete. The trajectories of (1) and (2) through x coincide for all $x \in D$, and thus the phase portraits of (1) and (2) are the same.

(2) Let M be a differentiable manifold of class C^∞ . A vector field of class C^r on M ($1 \leq r \leq \infty$) gives rise to an autonomous system

of ordinary differential equations in a coordinate neighborhood of each point of M , and it generates a local one-parameter group of local transformations of class C^r . If M is compact, this local one-parameter group of local transformations is uniquely extended to a one-parameter group of transformations of class C^r (\rightarrow 105 Differentiable Manifolds). Thus a vector field of class C^r on M generates a unique C^r -flow on M if M is compact. Therefore, sometimes we identify a vector field of class C^r with the C^r -flow generated by it.

(3) Let (M, φ) be a C^r -flow. Then $\varphi_1: M \rightarrow M$ is a C^r -diffeomorphism, which we call the **time-one mapping (time-one map)** of (M, φ) . Thus every C^r -flow (M, φ) induces a C^r -diffeomorphism as a time-one mapping. But the set of C^1 -diffeomorphisms that are time-one mappings of C^1 -flows is of the first category in the space of all C^1 -diffeomorphisms with C^1 topology (J. Palis). Thus most C^1 -diffeomorphisms are not expressed as time-one mappings of C^1 -flows.

(4) Let M be a compact differentiable manifold of class C^∞ with dimension m and (M, φ) a C^r -flow ($1 \leq r \leq \infty$). An $(m-1)$ -dimensional closed submanifold Σ of M is called a **cross section** of (M, φ) if the following conditions are satisfied: (i) For any $x \in X$, there exist $t_1 > 0$ and $t_2 < 0$ such that $\varphi_{t_1}(x), \varphi_{t_2}(x) \in \Sigma$; (ii) Every orbit intersects Σ transversally whenever it meets Σ . Let Σ be a cross-section of (M, φ) and $x \in \Sigma$. Let t_1 be the least positive number with $\varphi_{t_1}(x) \in \Sigma$. Such a t_1 exists for every $x \in \Sigma$, and $f: \Sigma \rightarrow \Sigma$ defined by $f(x) = \varphi_{t_1}(x)$, $x \in \Sigma$ is a C^r -diffeomorphism. We call this diffeomorphism f the **first-return mapping (first-return map)** or the **Poincaré mapping (Poincaré map)** for Σ .

(5) Let N be a compact differentiable manifold of class C^∞ and $f: N \rightarrow N$ a C^r -diffeomorphism ($1 \leq r \leq \infty$). Define an equivalence relation \sim on $N \times \mathbb{R}$ generated by $(x, t + 1) \sim (f(x), t)$ for $x \in X$, $t \in \mathbb{R}$. Then the quotient space $N(f) = N \times \mathbb{R} / \sim$ has a natural differentiable structure of class C^r , and a C^r -flow $(N(f), \psi)$ is defined by $\psi([x, t], s) = [x, t + s]$ for $x \in N$, $t, s \in \mathbb{R}$, where $[x, t] \in N(f)$ is the equivalence class of (x, t) . The flow $(N(f), \psi)$ thus obtained is called the **suspension** of f . The suspension $(N(f), \psi)$ has a cross section $\Sigma = \{[x, 0] | x \in N\}$, and the Poincaré mapping for Σ is C^r -conjugate to f . Conversely, if (M, φ) has a cross section Σ and the Poincaré mapping for Σ is $f: \Sigma \rightarrow \Sigma$, then the suspension $(\Sigma(f), \psi)$ is C^r -equivalent to (M, φ) .

(6) Let U be an open set of \mathbb{R}^n and $f: U \rightarrow \mathbb{R}^n$ a continuous mapping. Consider the difference equation

$$x_{m+1} = f(x_m), \quad x_m \in U. \quad (3)$$

For each $x \in U$, let $\varphi(x, m)$ be the solution of (3)

with $\varphi(x, 0) = x$. If $f(U) \subset U$, then $\varphi(x, m)$ exists for all $m \in \mathbf{Z}_+$ and $x \in U$, and φ defines a discrete semiflow on U . If $f: U \rightarrow U$ is a homeomorphism, then $\varphi(x, m)$ exists for all $m \in \mathbf{Z}$ and $x \in U$, and φ defines a discrete flow on U (\rightarrow 104 Difference Equations).

D. Basic Concepts

For simplicity we assume that the phase spaces of dynamical systems are metric spaces, and we denote their metrics by d .

(1) Let (X, φ) be a continuous flow on a metric space X and $x \in X$. A subset A of X is called **invariant** if $\varphi_t(A) \subset A$ for all $t \in \mathbf{R}$. A subset of X is invariant if and only if it is a union of orbits. A subset A of X is **positively** (resp. **negatively**) **invariant** if $\varphi_t(A) \subset A$ for all $t \geq 0$ (resp. $t \leq 0$). The subset $C_+(x) = \{\varphi^x(t) | t \geq 0\}$ (resp. $C_-(x) = \{\varphi^x(t) | t \leq 0\}$) is called the **positive** (resp. **negative**) **semiorbit** or the **positive** (resp. **negative**) **half-trajectory** starting from x . A subset is positively (resp. negatively) invariant if and only if it is a union of positive (resp. negative) semiorbits. A subset is invariant if and only if it is both positively and negatively invariant.

The union and the intersection of invariant sets are invariant. If A is invariant, then its closure \bar{A} , its interior $\text{Int } A$, its boundary ∂A , and its complement $A^c = X - A$ are invariant. If A is invariant, then $\varphi(A \times \mathbf{R}) \subset A$ and the restriction mapping $\varphi|_A \times \mathbf{R}$ defines a continuous flow $(A, \varphi|_A \times \mathbf{R})$ on A . The flow thus obtained is called the **restriction** of (X, φ) on A .

(2) A point $x \in X$ is a **singular point**, an **equilibrium point**, a **critical point**, a **rest point**, or a **fixed point** if $C(x) = \{x\}$ (\rightarrow 290 Nonlinear Oscillation). A point is **regular** or **nonsingular** if it is not a singular point. The set of all singular points is a closed invariant set, and the set of all nonsingular points is an open invariant set. If A is a positively invariant set which is homeomorphic to the closed unit ball in \mathbf{R}^n , then there exists a singular point in A (*Brouwer fixed-point theorem).

A point $x \in X$ is **periodic** if there exists a $T \neq 0$ such that

$$\varphi(x, t + T) = \varphi(x, t) \quad (4)$$

holds for all $t \in \mathbf{R}$. If x is periodic, the motion φ^x and the orbit $C(x)$ are said to be **periodic**. A point x is periodic if and only if there exists a $T \neq 0$ with $\varphi(x, T) = x$. A singular point is periodic. For a periodic point x , a number T satisfying (4) is called a **period** of x . If x is nonsingular and periodic, then there exists a smallest positive period T_0 of x , and any period is an integral multiple of T_0 . An orbit of a nonsingular periodic point is called a **closed**

orbit. If C is a closed orbit, then all points of C are nonsingular periodic points, and their smallest positive periods coincide. A closed orbit is compact.

(3) Let x be a point of X . A point $y \in X$ is called an ω -**limit** (resp. α -**limit**) **point** or a **positive** (resp. **negative**) **limit point** of x if there exists a sequence $\{t_n\}$ of real numbers such that (i) $t_n \rightarrow \infty$ (resp. $t_n \rightarrow -\infty$) as $n \rightarrow \infty$ and (ii) $\varphi(x, t_n) \rightarrow y$ as $n \rightarrow \infty$. The set of all ω -limit (resp. α -limit) points of x is denoted by $\omega(x)$ (resp. $\alpha(x)$) and is called the ω -**limit** (resp. α -**limit**) **set** of x . For each $x \in X$, $\omega(x)$ and $\alpha(x)$ are closed invariant sets, and the following equalities hold: $C_+(x) = C_+(x) \cup \omega(x)$, $C_-(x) = C_-(x) \cup \alpha(x)$, and $C(x) = C(x) \cup \alpha(x) \cup \omega(x)$. If x is a periodic point, $\bar{C}(x) = C(x) = \alpha(x) = \omega(x)$. If C is a closed orbit, then $\bar{C} = C = C(x) = \alpha(x) = \omega(x)$ for all $x \in C$. If A is a compact invariant set and $x \in A$, then $\alpha(x)$ and $\omega(x)$ are nonempty.

Assume that X is locally compact and $x \in X$. Then $\omega(x)$ is connected if it is compact, and none of the connected components of $\omega(x)$ is compact if $\omega(x)$ is not compact.

(4) Let x be a point of X . Let $J_+(x)$ (resp. $J_-(x)$) be the set of all points y satisfying the following condition: There exist a sequence $\{t_n\}$ of numbers and a sequence $\{x_n\}$ of points in X such that (i) $t_n \rightarrow \infty$ (resp. $t_n \rightarrow -\infty$) as $n \rightarrow \infty$, (ii) $x_n \rightarrow x$ as $n \rightarrow \infty$, and (iii) $\varphi(x_n, t_n) \rightarrow y$ as $n \rightarrow \infty$. The set $J_+(x)$ (resp. $J_-(x)$) is a closed invariant set containing $\omega(x)$ (resp. $\alpha(x)$), called the **first positive** (resp. **negative**) **prolongational limit set** of x . If X is locally compact, then $J_+(x)$ is connected if it is compact, and none of the connected components of $J_+(x)$ is compact when $J_+(x)$ is not compact.

Notions of higher prolongations have been defined and investigated by T. Ura, J. Auslander, and P. Seibert.

(5) For a discrete flow, we can similarly define basic notions such as an invariant set, fixed point, periodic point, and so on defined in Sections D(1)–D(4) by replacing \mathbf{R} by \mathbf{Z} . The propositions and theorems stated above hold for discrete flows, except those concerning connectedness.

E. Recursive Concepts and Dispersive Concepts

(1) Let (X, φ) be a flow on a metric space X . Let $x \in X$ be a point such that there exist a neighborhood U of x and a $T > 0$ satisfying the condition $U \cap \varphi_t(U) = \emptyset$ for $t \geq T$. Then x is called **wandering**. The set of all wandering points is an open invariant set. A point is **nonwandering** if it is not wandering. The set Ω of all nonwandering points is a closed invariant set and is called the **nonwandering set**.

The following conditions are equivalent: (i) x is nonwandering, (ii) $x \in J_+(x)$, (iii) $x \in J_-(x)$. The nonwandering set Ω contains all singular points, closed orbits, and $\omega(x)$ and $\alpha(x)$ for all $x \in X$.

(2) A point $x \in X$ is **positively** (resp. **negatively**) **Poisson stable** if $x \in \omega(x)$ (resp. $x \in \alpha(x)$). It is **Poisson stable** if it is both positively and negatively Poisson stable. A positively Poisson stable point and a negatively Poisson stable point are nonwandering. The following conditions are equivalent: (i) x is positively Poisson stable, (ii) $\overline{C_+(x)} = \omega(x)$, (iii) $C(x) \subset \omega(x)$, (iv) for any neighborhood U of x and $T > 0$, there exists a $t > T$ with $\varphi(x, t) \in U$. If X is \dagger complete and $x \in X$ is positively Poisson stable but not periodic, then $\omega(x) - C(x)$ is \dagger dense in $\omega(x)$. This implies that if X is complete, then $C(x)$ is periodic if and only if $C(x) = \omega(x)$.

(3) If all the points of the phase space X are wandering, then (X, φ) is called **completely unstable**. If all the points of X are nonwandering, then (X, φ) is called **regionally recurrent**. If A is an invariant set such that the restriction of (X, φ) on A is regionally recurrent, then (X, φ) is said to be **regionally recurrent on A** . If (X, φ) is regionally recurrent and X is locally compact, then the set of all Poisson stable points is dense in X .

For a given flow (X, φ) , we obtain a sequence of invariant sets $\{\Omega_n\}$ and a sequence of restriction flows $\{(\Omega_n, \varphi_n)\}$ such that (i) $(\Omega_0, \varphi_0) = (X, \varphi)$; (ii) Ω_{n+1} is the nonwandering set of (Ω_n, φ_n) , $n \geq 0$; and (iii) $(\Omega_{n+1}, \varphi_{n+1})$ is the restriction of (X, φ) on Ω_{n+1} . Then $X = \Omega_0 \supset \Omega_1 \supset \dots \supset \Omega_n \supset \dots$. Put $\Omega_\omega = \bigcap_n \Omega_n$. Then Ω_ω is an invariant set of (X, φ) , and we denote the restriction of (X, φ) on Ω_ω by $(\Omega_\omega, \varphi_\omega)$. Starting from $(\Omega_\omega, \varphi_\omega)$, we obtain similarly a sequence of invariant sets $\{\Omega_{\omega+n}\}$ and a sequence of flows $\{(\Omega_{\omega+n}, \varphi_{\omega+n})\}$. If we obtain an ordinal number γ such that $\Omega_\gamma = \Omega_{\gamma+1} \neq \emptyset$ by continuing this process, then we call Ω_γ the set of **central motions**. In this case, the flow $(\Omega_\gamma, \varphi_\gamma)$ is regionally recurrent, and every invariant subset of X on which (X, φ) is regionally recurrent is contained in Ω_γ . When X is \dagger separable and Ω_1 is compact and nonempty, then the minimum of such ordinal γ is an \dagger ordinal of at most the second number class.

(4) Let \mathcal{R} be the set of all $x \in X$ satisfying the following condition: For each ε , $T > 0$, there exist a sequence $\{x_0 = x, x_1, \dots, x_k = x\}$ of points in X and a sequence $\{t_0, t_1, \dots, t_{k-1}\}$ of numbers such that $t_i \geq T$ and $d(\varphi_{t_i}(x_i), x_{i+1}) < \varepsilon$ for $0 \leq i \leq k-1$. The set \mathcal{R} is a closed invariant set containing the nonwandering set Ω and is called the **chain recurrent set**. If $X = \mathcal{R}$, then (X, φ) is called **chain recurrent**. The restriction $(\mathcal{R}, \varphi|_{\mathcal{R} \times \mathbf{R}})$ of (X, φ) on \mathcal{R} is chain recurrent (C. C. Conley).

(5) A nonempty closed invariant set is called a **minimal set** if none of its nonempty proper closed subsets is invariant. For a nonempty compact subset A of X , the following conditions are equivalent: (i) A is minimal, (ii) $\overline{C(x)} = A$ for all $x \in A$, (iii) $\overline{C_+(x)} = A$ for all $x \in A$, (iv) $\overline{C_-(x)} = A$ for all $x \in A$, (v) $\omega(x) = A$ for all $x \in A$, (vi) $\alpha(x) = A$ for all $x \in A$. A point $x \in X$ is **positively** (resp. **negatively**) **Lagrange stable** if $\overline{C_+(x)}$ (resp. $\overline{C_-(x)}$) is compact. If $\overline{C(x)}$ is compact, then x is called **Lagrange stable**. Every nonempty compact invariant set contains a compact minimal set. If $x \in X$ is positively (resp. negatively) Lagrange stable, then $\omega(x)$ (resp. $\alpha(x)$) contains a compact minimal set.

A closed invariant set A is called a **saddle set** if there exists a neighborhood U of A such that every neighborhood of A contains a point x such that $C_+(x) \not\subset U$ and $C_-(x) \not\subset U$. Otherwise A is called a **nonsaddle set**. For a point x of X , let $E(x)$ be the subset of X consisting of the points y satisfying the following condition: There exist a sequence $\{x_n\}$ of points in X and two sequences $\{t_n\}$, $\{s_n\}$ of numbers such that (i) $x_n \rightarrow x$, $t_n \rightarrow \infty$, $s_n \rightarrow -\infty$ as $n \rightarrow \infty$ and (ii) $\varphi(x_n, t_n) \rightarrow y$, $\varphi(x_n, s_n) \rightarrow y$ as $n \rightarrow \infty$. For a subset B of X , put $E(B) = \bigcup_{x \in B} E(x)$. Let $\{S_\alpha\}$ be the family of all saddle minimal sets and $\{F_\beta\}$ the family of all nonsaddle minimal sets. If the phase space X is compact, then the nonwandering set $\Omega = (\bigcup_\beta F_\beta) \cup (\bigcup_\alpha E(S_\alpha)) = (\bigcup_\beta F_\beta) \cup E(\bigcup_\alpha S_\alpha)$ (T. Saito).

(6) A point $x \in X$ is said to be **recurrent** if for any $\varepsilon > 0$ there exists a $T > 0$ such that the ε -neighborhood U of $\varphi^x([t, t+T])$ contains $C(x)$ for all $t \in \mathbf{R}$. If x is recurrent, the motion φ^x and the orbit $C(x)$ are said to be **recurrent**. Every orbit of a compact minimal set is recurrent, and if the phase space is complete, then the closure of a recurrent orbit is a compact minimal set (Birkhoff).

A set D of real numbers is called **relatively dense** if there exists a $T > 0$ such that $D \cap (t, t+T) \neq \emptyset$ for all $t \in \mathbf{R}$. Assume that $x \in X$ is Lagrange stable, then x is recurrent if and only if for every $\varepsilon > 0$ the set $\{t \mid d(x, \varphi(x, t)) < \varepsilon\}$ is relatively dense.

(7) A flow (X, φ) is **dispersive** if for any $x, y \in X$ there exist neighborhoods U_x of x and U_y of y and a $T > 0$ such that $U_x \cap \varphi_t(U_y) = \emptyset$ for all $t \geq T$. The following conditions are equivalent: (i) (X, φ) is dispersive; (ii) For any $x, y \in X$, there exist neighborhoods U_x of x and U_y of y and a $T > 0$ such that $U_x \cap \varphi_t(U_y) = \emptyset$ if $|t| \geq T$; (iii) $J_+(x) = \emptyset$ for all $x \in X$.

A flow (X, φ) is **parallelizable** if there exist a subset S of X and a homeomorphism $h: X \rightarrow S \times \mathbf{R}$ such that (i) $\varphi(S \times \mathbf{R}) = X$ and (ii) $h \circ \varphi(x, t) = (x, t)$ for all $x \in S$ and $t \in \mathbf{R}$. A flow (X, φ) is parallelizable if and only if there is a subset S of X satisfying the following con-

ditions: (i) For each $x \in X$ there exists a unique $\tau(x) \in \mathbf{R}$ with $\varphi(x, \tau(x)) \in S$; (ii) $\tau: X \rightarrow \mathbf{R}$ is continuous. A parallelizable flow is dispersive. A flow on a locally compact separable metric space is parallelizable if and only if it is dispersive (V. V. Nemytskiĭ, V. V. Stepanov).

An open set U of X is called a **tube** if there exist a $T > 0$ and a subset S of U satisfying the following conditions: (i) $\varphi(S \times (-T, T)) = U$, (ii) For each $x \in U$ there is a unique $\tau(x) \in \mathbf{R}$ with $|\tau(x)| < T$ such that $\varphi(x, \tau(x)) \in S$. The set S in the above definition is called a **local section**. If $x \in X$ is a regular point, then there exists a tube containing x (M. Bebutov, H. Whitney). The notion of a local section is a generalization of Poincaré's "surface sans contact" [1] or Birkhoff's "surface of section" [4], and it is related to the notion of the cross section.

(8) For discrete flows (homeomorphisms), basic notions, such as a nonwandering set, Poisson stability, regional recurrence, central motion, and minimal set, are defined similarly, and many of the propositions and theorems in Sections E(1)–E(5) hold similarly for discrete flows (homeomorphisms).

F. Stability

(1) Let (X, φ) be a continuous flow on a metric space X . A point $x \in X$ is called **orbitally stable** if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $C_+(y)$ is contained in the ε -neighborhood of $C_+(x)$ for all y with $d(x, y) < \delta$ (\rightarrow 394 Stability). A nonempty set A is called **stable** if every neighborhood of A contains a positively invariant neighborhood of A . If A is compact (in particular, if A is a periodic orbit), then orbital stability and stability for A are equivalent. A nonempty set A is called **asymptotically stable** if A is stable and there exists a neighborhood V of A such that $\omega(x) \subset A$ for all $x \in V$. If A is stable and $\omega(x) \subset A$ for all $x \in X$, then A is called **globally asymptotically stable**. A point $x \in X$ is **Lyapunov stable** if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(\varphi_t(x), \varphi_t(y)) < \varepsilon$ for all $t \geq 0$ and y with $d(x, y) < \delta$. Lyapunov stability implies orbital stability. A point x is **uniformly Lyapunov stable** if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that for $z \in C(x)$ and y with $d(y, z) < \delta$ we have $d(\varphi_t(y), \varphi_t(z)) < \varepsilon$ for all $t \geq 0$. Uniform Lyapunov stability implies Lyapunov stability. For a singular point, the notions of uniform Lyapunov stability, Lyapunov stability, orbital stability, and stability are equivalent. Assume that the phase space X is locally compact and A is a nonempty compact subset of X . Then A is asymptotically stable if and only if there exist a neighborhood N of A and a continuous real-valued function L on N such that (i) $L(x) = 0$ if $x \in A$ and $L(x) > 0$ if $x \notin A$; (ii) $L(\varphi(x, t)) < L(x)$ if $x \notin A$, $t > 0$, and $\varphi(\{x\} \times [0, t]) \subset N$.

Such a function L is called a **Lyapunov function** for A (\rightarrow 394 Stability). Assume that X is locally compact and A is nonempty, stable, and invariant. Then A is asymptotically stable if and only if there exists a neighborhood U of A such that any invariant set in U is necessarily contained in A (Ura). Let A be a nonempty closed invariant set. A is called an **attractor** if it has an open neighborhood U satisfying the following conditions: (i) U is positively invariant; (ii) For each open neighborhood V of A , there is a $T > 0$ such that $\varphi_t(U) \subset V$ for all $t \geq T$. Condition (ii) implies that $\bigcap_{t \geq 0} \varphi_t(U) = A$ and $\omega(x) \subset A$ for all $x \in U$. Thus an attractor is asymptotically stable. If A is an attractor, the **basin** of A is the union of all open neighborhoods of A satisfying (i) and (ii).

(2) Assume that the phase space X is complete. A motion φ^x ($x \in X$) is called **almost periodic** if for any $\varepsilon > 0$ there exists a relatively dense subset $\{\tau_n\}$ in \mathbf{R} such that $d(\varphi^x(t), \varphi^x(t + \tau_n)) < \varepsilon$ for all $t \in \mathbf{R}$ and τ_n (\rightarrow 18 Almost Periodic Functions). If φ^x is almost periodic, then φ^y is almost periodic for all $y \in C(x)$. If A is a compact minimal set and if φ^x is almost periodic for some x in A , then every motion φ^y , $y \in A$, is almost periodic. An almost periodic motion is recurrent. The converse is not true. But if x is recurrent and Lyapunov stable in $C(x)$ (i.e., in the restriction flow $(C(x), \varphi|_{C(x) \times \mathbf{R}})$), then φ^x is almost periodic. If x is uniformly Lyapunov stable in $C(x)$ and negatively Lagrange stable, then φ^x is almost periodic (A. A. Markov).

G. Singular Points and Closed Orbits

In this section we assume that the phase space is a * paracompact differentiable manifold of class C^∞ with metric d .

(1) Let E be a finite-dimensional real vector space and $L: E \rightarrow E$ a linear automorphism. L is called **hyperbolic** if it has no * eigenvalues of absolute value 1. If $L: E \rightarrow E$ is hyperbolic, there are unique vector subspaces E^s and E^u satisfying the following conditions: (i) $E = E^s \oplus E^u$, (ii) $L(E^s) = E^s$ and $L(E^u) = E^u$, (iii) if $\|\cdot\|$ is a * norm on E , then there exist constants $c > 0$ and $0 < \lambda < 1$ such that, for any positive integer m , $\|L^m(v)\| \leq c\lambda^m\|v\|$ when $v \in E^s$ and $\|L^{-m}(v)\| \leq c\lambda^m\|v\|$ when $v \in E^u$. The zero 0 of E is the only fixed point of a hyperbolic linear automorphism. Put $s = \dim E^s$ and $u = \dim E^u$. Then $s + u = \dim E$, and s (resp. u) is the number of eigenvalues of absolute value < 1 (resp. > 1) counted with multiplicity. A topological conjugacy class of a hyperbolic linear automorphism $L: E \rightarrow E$ is determined by s , u , and the signs of $\det(L|_{E^s})$ and $\det(L|_{E^u})$, where $\det(L|_{E^s})$ is the * determinant of the restriction

$L|E^\sigma: E^\sigma \rightarrow E^\sigma$ of L on E^σ ($\sigma = s, u$). Further investigations of topological classification of linear automorphisms have been carried out by N. H. Kuiper and J. W. Robbin.

(2) Let $f: \mathbf{R}^n \rightarrow \mathbf{R}^n$ be a C^r -diffeomorphism ($1 \leq r \leq \infty$). Assume that the origin $0 \in \mathbf{R}^n$ is a fixed point of f . Then the t differential $df_0: \mathbf{R}^n \rightarrow \mathbf{R}^n$ of f at 0 is a linear automorphism. It is given by $df_0(x) = J_0(f)x$, $x \in \mathbf{R}^n$, where $J_0(f)$ is the t Jacobian matrix of f at 0 and x is expressed as a column vector. If df_0 is hyperbolic, then f is topologically conjugate to df_0 in a sufficiently small neighborhood of 0 (P. Hartman). Assume that f is of class C^∞ , and let $\lambda_1, \dots, \lambda_n$ (possibly repeated) be the eigenvalues of df_0 . If $\lambda_i \neq \lambda_i^m$ for all $1 \leq i \leq n$ and for all nonnegative integers m_1, \dots, m_n with $2 \leq m_1 + \dots + m_n$, then in a sufficiently small neighborhood of 0, f is C^∞ -conjugate to df_0 (S. Sternberg).

(3) Let $f: M \rightarrow M$ be a C^r -diffeomorphism ($1 \leq r \leq \infty$) of a differentiable manifold M of class C^∞ . Let $p \in M$ be a fixed point of f . Then the t differential df_p of f at p is a linear automorphism of the t tangent space $T_p(M)$ of M at p . If df_p is hyperbolic, then p is called a **hyperbolic fixed point** of f . Let p be a hyperbolic fixed point of f and $T_p(M) = E^s \oplus E^u$ be the direct sum decomposition with respect to df_p . Put $W^s(p) = \{x \in M \mid f^n(x) \rightarrow p \text{ as } n \rightarrow \infty\}$ and $W^u(p) = \{x \in M \mid f^{-n}(x) \rightarrow p \text{ as } n \rightarrow \infty\}$. $W^s(p)$ and $W^u(p)$ are invariant sets and images of t injective immersions of class C^r of vector spaces E^s and E^u , respectively. At the point p , $W^s(p)$ and $W^u(p)$ are tangent to E^s and E^u , respectively [21]. We call $W^s(p)$ (resp. $W^u(p)$) the **stable** (resp. **unstable**) **manifold** of f at p . In a sufficiently small neighborhood of a hyperbolic fixed point p of f , f is topologically conjugate to its differential df_p . Therefore a hyperbolic fixed point is an **isolated fixed point** (i.e., there are no fixed points in its sufficiently small neighborhood except itself). A hyperbolic fixed point p is called a **source** if $\dim W^s(p) = 0$ and a **sink** if $\dim W^u(p) = 0$. Otherwise it is a **saddle point**. The number of topological conjugacy classes of hyperbolic fixed points on an n -dimensional manifold is 4^n .

Let $p \in M$ be a periodic point of f and m the smallest positive period of p . Replacing f by f^m , we obtain notions of hyperbolicity, stable manifold, and so on for the periodic point p . We also obtain propositions and theorems similar to those stated above for periodic points.

(4) Let A be a real $n \times n$ matrix. Consider the system of linear differential equations

$$dx/dt = Ax, \quad x \in \mathbf{R}^n. \quad (5)$$

This equation generates a C^∞ -flow (\mathbf{R}^n, φ) , and $\varphi_t: \mathbf{R}^n \rightarrow \mathbf{R}^n$ is given by $\varphi_t(x) = e^{tA}x$, where e^{tA} is

the t exponential of the matrix tA . Thus φ_t is a linear automorphism for all $t \in \mathbf{R}$. The origin $0 \in \mathbf{R}^n$ is a singular point of (\mathbf{R}^n, φ) . If none of the real parts of the eigenvalues of A are zero, we call $0 \in \mathbf{R}^n$ a **hyperbolic singular point** of (5). If $0 \in \mathbf{R}^n$ is a hyperbolic singular point of (5), then there exist two vector subspaces E^s and E^u of \mathbf{R}^n satisfying the following conditions: (i) $\mathbf{R}^n = E^s \oplus E^u$; (ii) $A(E^s) = E^s$ and $A(E^u) = E^u$; (iii) E^s and E^u are invariant sets of (\mathbf{R}^n, φ) ; (iv) there exist positive constants c and λ such that for all $t \geq 0$, $\|\varphi_t(x)\| \leq ce^{-\lambda t} \|x\|$ when $x \in E^s$ and $\|\varphi_{-t}(x)\| \leq ce^{-\lambda t} \|x\|$ when $x \in E^u$, where $\|\cdot\|$ is the norm on \mathbf{R}^n . The origin 0 is a hyperbolic singular point of (5) if and only if the time-one mapping $\varphi_1 = e^A: \mathbf{R}^n \rightarrow \mathbf{R}^n$ of φ is a hyperbolic linear automorphism. If 0 is a hyperbolic singular point of (5), the above direct sum decomposition $\mathbf{R}^n = E^s \oplus E^u$ coincides with the one with respect to φ_1 . Put $s = \dim E^s$ and $u = \dim E^u$. Then $s + u = n$, and s (resp. u) is the number of all eigenvalues of A with negative (resp. positive) real parts counted with multiplicity. Also we obtain the following properties: (i) $x \in E^s$ (resp. E^u) $\Leftrightarrow \varphi_t(x) \rightarrow 0$ as $t \rightarrow \infty$ (resp. $t \rightarrow -\infty$), (ii) $x \notin E^s$ (resp. E^u) $\Rightarrow \|\varphi_t(x)\| \rightarrow \infty$ as $t \rightarrow \infty$ (resp. $t \rightarrow -\infty$), (iii) the origin 0 is the only singular point.

Let A and B be $n \times n$ real matrices. Let $(\mathbf{R}^n, \varphi_A)$ and $(\mathbf{R}^n, \varphi_B)$ be the flows determined by the equations $dx/dt = Ax$ and $dx/dt = Bx$, respectively. Assume that $0 \in \mathbf{R}^n$ is a hyperbolic singular point for both of the equations. Let s and u (resp. s' and u') be the integers defined for φ_A (resp. φ_B) in the above paragraph. Then the following conditions are equivalent: (i) $(\mathbf{R}^n, \varphi_A)$ and $(\mathbf{R}^n, \varphi_B)$ are flow equivalent, (ii) $(\mathbf{R}^n, \varphi_A)$ and $(\mathbf{R}^n, \varphi_B)$ are topologically equivalent, (iii) $s = s'$, (iv) $u = u'$, (v) $(\varphi_A)_1$ and $(\varphi_B)_1$ are topologically conjugate. Further investigations of the phase portrait of the equation $dx/dt = Ax$ without the assumption of hyperbolicity were done by Kuiper.

(5) Let D be an open set of \mathbf{R}^n and $f: D \rightarrow \mathbf{R}^n$ a continuous mapping. Consider the system of differential equations (1), which we write down again here:

$$dx/dt = f(x), \quad x \in D. \quad (6)$$

A point $p \in D$ is a **singular point** of (6) if the trajectory through p consists of a single point p . If (6) generates a flow (D, φ) , then p is a singular point of (6) if and only if p is a singular point of (D, φ) . A point p is a singular point of (6) if and only if $f(p) = 0$.

Suppose further that f is a C^r -mapping ($1 \leq r \leq \infty$). A singular point p of (6) is called **hyperbolic** if none of the real parts of the eigenvalues of the Jacobian matrix $J_p(f)$ of f at p is zero. Let p be a singular point of (6). We assume for simplicity that $p = 0$ and (6) generates

a flow (D, φ) . Denote the Jacobian matrix of f at 0 by A , and let $(\mathbf{R}^n, \varphi_A)$ be the flow generated by the equation $dx/dt = Ax$. If 0 is a hyperbolic singular point, then in a sufficiently small neighborhood of 0, the flows (D, φ) and $(\mathbf{R}^n, \varphi_A)$ are flow equivalent and hence topologically equivalent (Hartman, D. M. Grobman). Let $\lambda_1, \dots, \lambda_n$ (possibly repeated) be the eigenvalues of A . If f is of class C^∞ and $\lambda_i \neq m_1 \lambda_1 + \dots + m_n \lambda_n$ for all $1 \leq i \leq n$ and for all nonnegative integers m_1, \dots, m_n with $2 \leq m_1 + \dots + m_n$, then in a sufficiently small neighborhood of 0, two flows (D, φ) and $(\mathbf{R}^n, \varphi_A)$ are C^∞ -equivalent (Sternberg).

(6) Let M be a paracompact differentiable manifold of class C^∞ and V a vector field of class C^r ($1 \leq r \leq \infty$) on M . For simplicity we assume that V generates a C^r -flow (M, φ) . For each point p of M , take a coordinate neighborhood (U, α) of class C^∞ around p , where U is an open neighborhood of p in M and $\alpha: U \rightarrow \mathbf{R}^n$ is a homeomorphism onto an open set D in \mathbf{R}^n . Using (U, α) , we can express the vector field V as a system of ordinary differential equations of the form (6) in D . The eigenvalues $\lambda_1, \dots, \lambda_n$ (possibly repeated) of the Jacobian matrix of f at $\alpha(p)$ are independent of the choice of local coordinates (U, α) around p . Thus $\lambda_1, \dots, \lambda_n$ are invariants of the C^r -equivalence, but they are not invariants of the topological equivalence.

A singular point p of a flow (M, φ) (or a vector field V) is called **hyperbolic** if $\alpha(p)$ is a hyperbolic singular point of the corresponding equation (6) (i.e., none of the real parts of the above eigenvalues $\lambda_1, \dots, \lambda_n$ are zero). A singular point of (M, φ) is hyperbolic if and only if it is a hyperbolic fixed point of the time-one mapping φ_1 of φ . A hyperbolic singular point p is an isolated singular point.

Let p be a hyperbolic singular point of (M, φ) and $T_p(M) = E^s \oplus E^u$ the direct sum decomposition with respect to φ_1 . Put $s = \dim E^s$ and $u = \dim E^u$. Then s (resp. u) is the number of all eigenvalues λ_i of negative (resp. positive) real parts. Put $W^s(p) = \{x \in M \mid \varphi_t(x) \rightarrow p \text{ as } t \rightarrow \infty\}$ and $W^u(p) = \{x \in M \mid \varphi_{-t}(x) \rightarrow p \text{ as } t \rightarrow \infty\}$. $W^s(p)$ and $W^u(p)$ are invariant sets and images of injective immersions of class C^r of vector spaces E^s and E^u , respectively. At the point p , $W^s(p)$ and $W^u(p)$ are tangent to E^s and E^u , respectively. We call $W^s(p)$ (resp. $W^u(p)$) the **stable** (resp. **unstable**) manifold of f at p . The stable manifold and the unstable manifold of (M, φ) at a hyperbolic singular point p coincide with those of the time-one mapping φ_1 at p . A hyperbolic singular point p is called a **source** if $\dim W^s(p) = 0$ and a **sink** if $\dim W^u(p) = 0$. Otherwise it is called a **saddle point**. If p is a source (resp. a sink), then $W^u(p)$ (resp. $W^s(p)$) is a neighborhood of p . A singular point p is a source (resp. a sink) if and

only if all real parts of the corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ are positive (resp. negative).

(7) Let V be a vector field of class C^r on a C^∞ -differentiable manifold M of dimension $n \geq 2$ and (M, φ) the C^r -flow generated by V . Let C be a closed orbit of (M, φ) , p a point of C , and $T > 0$ its smallest positive period. Then p is a fixed point of the C^r -diffeomorphism $\varphi_T: M \rightarrow M$. Let $V(p) \in T_p(M)$ be the value of V at p and $(d\varphi_T)_p: T_p(M) \rightarrow T_p(M)$ the differential of φ_T at p . Then $(d\varphi_T)_p(V(p)) = V(p)$, and $V(p) \neq 0$. Therefore 1 is an eigenvalue of $(d\varphi_T)_p$. The other eigenvalues $\lambda_1, \dots, \lambda_{n-1}$ (possibly repeated) of $(d\varphi_T)_p$ do not depend on a choice of $p \in C$ and are called the **characteristic multipliers** of C . If none of the characteristic multipliers of C is of absolute value 1, we call C a **hyperbolic closed orbit**. If C is a hyperbolic closed orbit, then there exist vector subspaces E_p^s and E_p^u for each $p \in C$ satisfying the following conditions: (i) $T_p(M) = L(V(p)) \oplus E_p^s \oplus E_p^u$, where $L(V(p))$ is the 1-dimensional subspace generated by $V(p)$; (ii) $(d\varphi_t)_p(E_p^s) = E_{\varphi_t(p)}^s$ ($\sigma = s, u$) for all $t \in \mathbf{R}$, where $(d\varphi_t)_p: T_p(M) \rightarrow T_{\varphi_t(p)}(M)$. In particular, $(d\varphi_T)_p(E_p^s) = E_p^s$ ($\sigma = s, u$); (iii) $\dim E_p^s$ (resp. $\dim E_p^u$) is independent of $p \in C$ and is equal to the number of λ_i of absolute value < 1 (resp. > 1). Put $W^s(C) = \{x \in M \mid d(\varphi_t(x), C) \rightarrow 0 \text{ as } t \rightarrow \infty\}$ and $W^u(C) = \{x \in M \mid d(\varphi_{-t}(x), C) \rightarrow 0 \text{ as } t \rightarrow \infty\}$. Then $W^s(C)$ (resp. $W^u(C)$) is an invariant set and is an injectively immersed C^r -submanifold of M which is tangent to $L(V(p)) \oplus E_p^s$ (resp. $L(V(p)) \oplus E_p^u$) for each $p \in C$. $W^s(C)$ (resp. $W^u(C)$) is called the **stable** (resp. **unstable**) manifold for C .

Let p be a point of a closed orbit C . An embedded $(n-1)$ -dimensional disk D of class C^r in M containing p is called a **cross section for a closed orbit** C if V is transverse to D (i.e., $T_x(M) = L(V(x)) \oplus T_x(D)$ for each $x \in D$) and $D \cap C = \{p\}$. For a given cross section D for C , there exists a neighborhood U of p in D such that for any $x \in U$ there exists a $\tau(x) > 0$ such that $\varphi(x, \tau(x)) \in D$ and $\varphi(x, t) \notin D$ for $0 < t < \tau(x)$. A mapping $f: U \rightarrow D$, defined by $f(x) = \varphi(x, \tau(x))$, $x \in U$, is called a **Poincaré mapping** for D . The C^r -conjugacy class of the germ of a Poincaré mapping $f: V \rightarrow D$ at p is independent of the choice of $p \in C$ and the cross section D for the closed orbit C . The point p is a fixed point of f , and the eigenvalues of df_p coincide with the characteristic multipliers of C including multiplicity. Therefore C is hyperbolic if and only if $p \in C$ is a hyperbolic fixed point of f . In a sufficiently small neighborhood of C , the flow is C^r -equivalent to the suspension of f . The topological equivalence of the flow in a sufficiently small neighborhood of C is determined by the topological conjugacy class of f . The number of topological equivalence classes of hyperbolic closed orbits which appear in a

flow on an n -dimensional manifolds is $4(n-1)$ (M. C. Irwin).

(8) Let V be a vector field of class C^r ($1 \leq r \leq \infty$) on a compact C^∞ -manifold M . If M has a nonempty boundary, we assume that V points outward at all boundary points. Let p be an isolated singular point of V and denote by $i(p)$ the \dagger index of the vector field V at the singular point P . If all singular points are isolated, then the sum $\sum_p i(p)$ is independent of V and is equal to the \dagger Euler-Poincaré characteristic $\chi(M)$ of M (Poincaré, H. Hopf) (\rightarrow 153 Fixed-Point Theorems). A vector field (or a flow) is called **nonsingular** if it has no singular points. If $\chi(M) \neq 0$, then any vector field (and hence any flow) on M has a singular point. If $\chi(M) = 0$, then there exists a nonsingular vector field on M . If M is 2-dimensional and without boundary, then M admits a nonsingular vector field if and only if M is a \dagger torus or a \dagger Klein bottle.

There are many directions in which generalizations of the Poincaré-Hopf theorem can be made. For example, an index of an isolated closed orbit of a flow has been defined and investigated by F. B. Fuller.

The phase portrait near a singular point or a closed orbit which is not hyperbolic is complicated. Many results in general situations have been obtained for planar flows, but there are only a few for higher-dimensional flows.

H. Generic Properties and Structural Stability

(1) Let M be a compact C^∞ -manifold without boundary. The set $\mathcal{X}^r(M)$ of all vector fields of class C^r ($1 \leq r \leq \infty$) on M forms a real vector space in a natural way. We can give a norm $\|V\|_r$ (called a **C^r -norm**) for $V \in \mathcal{X}^r(M)$ using its expressions in a given suitable system of local coordinates of class C^∞ on M and their partial derivatives up to order r . By virtue of this norm, $\mathcal{X}^r(M)$ is a Banach space with the topology of uniform C^r -convergence (\rightarrow 168 Function Spaces). A subset of a topological space is called a **residual set** or a **Baire set** if it is the intersection of a countable number of dense open sets. A residual set in $\mathcal{X}^r(M)$ is a dense set. A proposition P concerning a vector field of class C^r is called **generic** if the set $\{V \in \mathcal{X}^r(M) | P(V)\}$ contains a residual set for any M . The following properties are generic properties: (i) All singular points and all closed orbits are hyperbolic [21, 22]; (ii) any stable manifold (of a hyperbolic singular point or a hyperbolic closed orbit) meets transversally with any unstable manifold (of a hyperbolic singular point or a hyperbolic closed orbit) at any point of their intersection [21, 22]; (iii) the set of all periodic points (i.e., the union of all

singular points and all closed orbits) is dense in the nonwandering set Ω for $V \in \mathcal{X}^1(M)$ [23]; (iv) there is no regular first integral for $v \in \mathcal{X}^1(M)$, where a **regular first integral** of a vector field V on M is a C^1 -function $f: M \rightarrow \mathbf{R}$ such that f is not constant on any open set of M and is constant along any orbit of (the flow generated by) V (R. C. Robinson).

Let V (resp. V') be an element of $\mathcal{X}^r(M)$ (resp. $\mathcal{X}^r(M')$), (M, φ) (resp. (M', φ')) the flow generated by V (resp. V'), and Ω (resp. Ω') the nonwandering set of (M, φ) (resp. (M', φ')). V and V' are called **topologically equivalent** (resp. **Ω -equivalent**) if (M, φ) and (M', φ') (resp. the restrictions $(\Omega, \varphi|_{\Omega \times \mathbf{R}})$ and $(\Omega', \varphi'|_{\Omega' \times \mathbf{R}})$) are topologically equivalent. A vector field $V \in \mathcal{X}^r(M)$ (or the flow generated by V) is called **C^r -structurally stable** (resp. **C^r - Ω -stable**) if there exists a neighborhood \mathcal{N} of V in $\mathcal{X}^r(M)$ such that any V' in \mathcal{N} is topologically equivalent (resp. Ω -equivalent) to V . If $V \in \mathcal{X}^r(M)$ is structurally stable (resp. Ω -stable), then the topological structure of the phase portrait of (the flow generated by) V in the whole space (resp. the nonwandering set) remains invariant under a sufficiently small C^r -perturbation of V . The generic properties (i)–(iv) above hold for C^1 -structurally stable vector fields (Markus, Thom, Peixoto, J. Arnaud). The generic properties (i) and (iii) above hold for C^1 - Ω -stable vector fields.

(2) Let M be a compact C^∞ -manifold without boundary. Let $F^r(M)$ be the set of all C^r -mappings of M into itself with the topology of uniform C^r -convergence. Let $\text{Diff}^r(M)$ be the subset of $F^r(M)$ consisting of all C^r -diffeomorphisms of M onto itself. Then $F^r(M)$ is a complete metric space, and $\text{Diff}^r(M)$ is open in $F^r(M)$. Thus $\text{Diff}^r(M)$ is a \dagger Baire space and also a topological group with respect to this topology. A proposition P concerning $f \in \text{Diff}^r(M)$ is called **generic** if the set $\{f \in \text{Diff}^r(M) | P(f)\}$ contains a residual set. The following properties are generic: (i) Every periodic point is hyperbolic [21]; (ii) for each pair of hyperbolic periodic points p and q , the stable manifold $W^s(p)$ intersects the unstable manifold $W^u(q)$ transversally [21]; (iii) for a C^1 -diffeomorphism $f \in \text{Diff}^1(M)$ the set of all periodic points is dense in the nonwandering set ([23], Robinson).

Let $f: M \rightarrow M$ and $g: N \rightarrow N$ be C^r -diffeomorphisms and $\Omega(f)$ and $\Omega(g)$ their nonwandering sets. f and g are called **Ω -conjugate** if $f|_{\Omega(f)}: \Omega(f) \rightarrow \Omega(f)$ and $g|_{\Omega(g)}: \Omega(g) \rightarrow \Omega(g)$ are topologically conjugate. A diffeomorphism $f \in \text{Diff}^r(M)$ is called **C^r -structurally stable** (resp. **C^r - Ω -stable**) if there exists a neighborhood \mathcal{N} of f in $\text{Diff}^r(M)$ such that any g in \mathcal{N} is topologically conjugate (resp. Ω -conjugate) to f . Generic properties (i)–

(iii) above hold for C^1 -structurally stable diffeomorphisms, and (i) and (iii) above hold for C^1 - Ω -stable diffeomorphisms.

I. Low-Dimensional Systems

(1) Let $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ be the unit circle in the complex plane \mathbb{C} and $p: \mathbb{R} \rightarrow S^1$ the \dagger covering projection defined by $p(x) = e^{2\pi i x}$, $x \in \mathbb{R}$. Let $f: S^1 \rightarrow S^1$ be an \dagger orientation-preserving homeomorphism. Then f can be lifted to a mapping $F: \mathbb{R} \rightarrow \mathbb{R}$ satisfying the following conditions: (i) $p \circ F = f \circ p$; (ii) F is monotone increasing; (iii) $F - 1_{\mathbb{R}}$ is a periodic function of period 1, where $1_{\mathbb{R}}$ is the identity mapping of \mathbb{R} . The limit $\rho(f) = \lim_{n \rightarrow \infty} F^n(x)/n$ exists for all $x \in \mathbb{R}$, and its \dagger residue class modulo \mathbb{Z} is independent of F and x . We call $\rho(f)$ the **rotation number** of f . Let $f, g: S^1 \rightarrow S^1$ be orientation-preserving homeomorphisms. If f and g are topologically conjugate by an orientation-preserving (resp. \dagger reversing) homeomorphism, then $\rho(f) \equiv \rho(g)$ (resp. $\rho(f) = -\rho(g)$) modulo \mathbb{Z} . An orientation preserving homeomorphism $f: S^1 \rightarrow S^1$ has a periodic point of the smallest positive period s if and only if $\rho(f) = r/s$, where r and s are relatively prime integers. In this case all periodic points are of the smallest positive period s . If $\rho(f)$ is irrational, then the ω -limit set $\omega(x)$ of $x \in S^1$ is independent of x , and $E = \omega(x)$ is either \dagger perfect and \dagger nowhere dense or the whole space S^1 . If $\rho(f)$ is irrational and $E = \omega(x) = S^1$, then f is called **transitive**. If f is transitive, then it is topologically conjugate to the rotation $r_{\rho(f)}: S^1 \rightarrow S^1$ defined by $r_{\rho(f)}(e^{2\pi i x}) = e^{2\pi i(x + \rho(f))}$, $x \in \mathbb{R}$. Let $f: S^1 \rightarrow S^1$ be of class C^1 with $\rho(f)$ irrational. If its derivative f' is of \dagger bounded variation, then f is transitive (A. Denjoy). In particular, if f is of class C^2 with $\rho(f)$ irrational, then f is topologically conjugate to the rotation $r_{\rho(f)}$. However, there are C^1 -diffeomorphisms of S^1 onto itself whose nonwandering sets are not the whole space. Those C^1 -diffeomorphisms are never topologically conjugate to C^2 -diffeomorphisms. M. R. Herman gave a sufficient condition for a diffeomorphism of S^1 onto itself to be differentiably conjugate to a rotation.

A C^r -diffeomorphism $f: S^1 \rightarrow S^1$ is structurally stable if and only if $\Omega(f)$ is finite (hence $\Omega(f)$ consists of a finite number of periodic points), and all periodic points are hyperbolic. The set of all structurally stable C^r -diffeomorphisms of S^1 onto itself is a dense open set in $\text{Diff}^r(S^1)$ (Peixoto).

(2) Let D be an open set in \mathbb{R}^2 and $f: D \rightarrow \mathbb{R}^2$ a continuous mapping. We assume that for each $x \in D$ there exists a unique nonextendable solution $\varphi(x, t)$ with the initial condition

$\varphi(x, 0) = x$ defined on (a_x, b_x) for the equation $dx/dt = f(x)$, $x \in D$. Suppose that for a given point $x \in D$ there exists a compact set K in D containing the positive semiorbit $C_+(x) = \{\varphi(x, t) \mid 0 \leq t < b_x\}$ of x . Then $b_x = \infty$. Further, we assume that the ω -limit set $\omega(x)$ of x contains no singular points. Then we have either (i) $C_+(x) = \omega(x)$ and it is a closed orbit, or (ii) $C_+(x) \neq \omega(x)$ and $\omega(x)$ is a closed orbit. In case (ii), $\omega(x)$ is a \dagger simple closed curve and $C_+(x)$ is a spiral which tends the closed orbit $\omega(x)$ (Poincaré, I. Bendixson). We call such a closed orbit $\omega(x)$ a **limit cycle**. Let f_1 and f_2 be polynomials of two variables and m the maximum of degrees of f_1 and f_2 . Let $f = (f_1, f_2): \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the mapping defined by f_1 and f_2 . The equation $dx/dt = f(x)$, $x \in \mathbb{R}^2$, defined by such an f is called a polynomial system of degree m . The following is Hilbert's 16th problem: Is there a number $N(m)$ depending only on m such that the number of limit cycles for any polynomial system of degree m is bounded by $N(m)$?

Let M be a closed (i.e., compact, without boundary) C^∞ -manifold of dimension 2 and (M, φ) a C^2 -flow on M . Then a minimal set of (M, φ) is either (i) a singular point, (ii) a closed orbit, or (iii) the whole space M . For case (iii), M is a torus (Poincaré, Denjoy, C. L. Siegel, A. J. Schwartz). Let T be a 2-dimensional torus and (T, φ) a C^r -flow. Suppose that (T, φ) has a cross section Σ which is C^1 -diffeomorphic to S^1 . Let $f: \Sigma \rightarrow \Sigma$ be the Poincaré mapping for Σ . Then (T, φ) has a closed orbit if and only if the rotation number $\rho(f)$ of f is rational. If $\rho(f)$ is irrational and (T, φ) is of class C^2 , then T is a minimal set.

Let M be an orientable 2-dimensional closed C^∞ -manifold and $V \in \mathcal{X}^1(M)$. Then V is structurally stable if and only if the following conditions are satisfied: (i) There are only a finite number of singular points, all hyperbolic; (ii) There are only a finite number of closed orbits, all hyperbolic; (iii) There are no orbits which connect saddle points; (iv) The $\alpha(x)$ and $\omega(x)$ for any $x \in M$ are singular points or closed orbits (Peixoto). The above theorem was first proved by Andronov and Pontryagin for analytic vector fields on a 2-dimensional disk which are transverse to the boundary of the disk at any boundary point. The set of all structurally stable vector fields of class C^1 on M is a dense open set in $\mathcal{X}^1(M)$ (Peixoto).

J. Axiom A Systems

In this section we assume that phase spaces are closed C^∞ -manifolds with metric d and $1 \leq r \leq \infty$ unless stated otherwise.

(1) A vector field $V \in \mathcal{X}^r(M)$ (resp. a C^r -flow (M, φ)) is called a **Morse-Smale vector field** (resp. a **Morse-Smale flow**) if the following conditions are satisfied: (i) The nonwandering set is the union of a finite number of singular points and a finite number of closed orbits; (ii) The singular points and closed orbits are all hyperbolic; (iii) The stable manifolds and the unstable manifolds of the singular points and closed orbits intersect each other transversely. If $\dim M = 2$, the Morse-Smale vector fields on M are exactly the structurally stable ones discussed in Section I(2). A Morse-Smale vector field is structurally stable [25]. The set of all Morse-Smale vector fields on M is open but not dense in $\mathcal{X}^r(M)$ if $\dim M > 2$ (Palis, Smale). However, it contains a dense open subset of the set of all \dagger gradient vector fields with respect to a given \dagger Riemannian metric (Smale). In particular, any closed manifold admits Morse-Smale vector fields and hence structurally stable vector fields. For a Morse-Smale vector field, the \dagger Morse inequalities hold as they do in the \dagger calculus of variations in the large [24].

(2) A vector field $V \in \mathcal{X}^r(M)$ (or the C^r -flow (M, φ) generated by V) is called an **Anosov vector field** (or an **Anosov flow**) if the following conditions are satisfied: (i) There is a direct sum decomposition $T_x(M) = L(V(x)) \oplus E_x^s \oplus E_x^u$ of the tangent space $T_x(M)$ for each $x \in M$ which depends continuously on $x \in M$; (ii) $(d\varphi_t)_x(E_x^s) = E_{\varphi_t(x)}^s$ and $(d\varphi_t)_x(E_x^u) = E_{\varphi_t(x)}^u$ for all $x \in M$ and $t \in \mathbf{R}$; (iii) There are a Riemannian metric on M and constants $c, \lambda > 0$ such that, for all $t > 0$ and $x \in M$, $\|(d\varphi_t)_x(w)\| \leq ce^{-\lambda t} \|w\|$ when $w \in E_x^s$, and $\|(d\varphi_{-t})_x(w)\| \leq ce^{-\lambda t} \|w\|$ when $w \in E_x^u$, where $\|\cdot\|$ is the norm induced by the Riemannian metric. The suspensions of Anosov diffeomorphisms and the \dagger geodesic flows on Riemannian manifolds of negative curvature are important examples of Anosov flows (J. Hadamard [8]). There are examples of Anosov flows other than the ones stated above (M. Handel and W. P. Thurston, J. Franks and R. F. Williams). The following have been proved by Anosov [8]: (i) An Anosov flow is structurally stable; (ii) there are countably many closed orbits for an Anosov flow; (iii) if there exists a **smooth invariant measure** (i.e., an \dagger invariant measure which has a \dagger smooth density with respect to the measure associated with the Riemannian metric), then the set of all closed orbits is dense in M . If we assume further that the flow is of class C^2 , then it is \dagger ergodic; (iv) the set of all Anosov vector fields is open in $\mathcal{X}^r(M)$; (v) $\{E_x^s\}, \{E_x^u\}$ ($x \in M$) define \dagger foliations on M , which are called **Anosov foliations**.

(3) A vector field $V \in \mathcal{X}^r(M)$ (or the C^r -flow (M, φ) generated by V) is called an **Axiom A**

vector field (or an **Axiom A flow**) if the following conditions are satisfied: A(a) The nonwandering set consists of a finite set F of singular points, all hyperbolic, and the closure Λ of the union of closed orbits, and $F \cap \Lambda = \emptyset$; A(b) The conditions (i)–(iii) in the definition of Anosov flow in which we replace the terms “ $x \in M$ ” by “ $x \in \Lambda$.” For each $x \in M$, put $W^s(x) = \{y \in M \mid d(\varphi_t(x), \varphi_t(y)) \rightarrow 0 \text{ as } t \rightarrow \infty\}$ and $W^u(x) = \{y \in M \mid d(\varphi_{-t}(x), \varphi_{-t}(y)) \rightarrow 0 \text{ as } t \rightarrow \infty\}$. We call $W^s(x)$ (resp. $W^u(x)$) the **stable** (resp. **unstable**) manifold of V at x . For a subset A of M , put $W^\sigma(A) = \bigcup_{x \in A} W^\sigma(x)$ ($\sigma = s, u$). For $x \in M$, put $W^{w\sigma}(x) = W^\sigma(C(x))$ ($\sigma = s, u$), where $C(x)$ is the orbit of x . If the flow satisfies Axiom A, then $W^s(x)$, $W^u(x)$, $W^{ws}(x)$, and $W^{wu}(x)$ are injectively immersed submanifolds for all $x \in M$ (M. W. Hirsch and C. C. Pugh). If $x \in M$ is a hyperbolic singular point, then $W^s(x)$ and $W^u(x)$ defined above coincide with those defined before. If $C(x)$ is a hyperbolic closed orbit, then $W^s(C(x))$ and $W^u(C(x))$ defined above coincide with those defined before.

For an Axiom A flow, there is a decomposition of the nonwandering set $\Omega = \Omega_1 \cup \dots \cup \Omega_m$ (disjoint union), where each Ω_i is closed, invariant, and **transitive** (i.e., has a dense orbit), and $M = \bigcup_{i=1}^m W^s(\Omega_i) = \bigcup_{i=1}^m W^u(\Omega_i)$ (disjoint union) [7]. This decomposition is called the **spectral decomposition** of Ω , and each Ω_i is called a **basic set**. Let $\Omega = \Omega_1 \cup \dots \cup \Omega_m$ be the spectral decomposition of the nonwandering set for an Axiom A flow. Denote $\Omega_i \leq \Omega_j$ if $W^s(\Omega_i) \cap W^u(\Omega_j) \neq \emptyset$. A sequence of basic sets $\Omega_{i_0}, \Omega_{i_1}, \dots, \Omega_{i_k}$ ($k > 1$) is called a **cycle** if $\Omega_{i_0} \leq \Omega_{i_1} \leq \dots \leq \Omega_{i_k}, \Omega_{i_0} = \Omega_{i_k}$, and otherwise $\Omega_{i_p} \not\leq \Omega_{i_q}$ for $p \neq q$. An Axiom A flow which has no cycles in the above sense is said to satisfy the **no cycle condition**. The **Ω -stability theorem**:

An Axiom A flow with the no cycle condition is Ω -stable [27]. **Ω -explosion**: If an Axiom A flow has a cycle, then it is not Ω -stable (Palis). An Axiom A flow is said to satisfy the **strong transversality condition** if $W^{ws}(x)$ and $W^{wu}(x)$ intersect transversely at any $x \in M$. An Axiom A flow with the strong transversality condition satisfies the no cycle condition [7]. The **structural stability theorem**: An Axiom A flow of class C^1 with the strong transversality condition is structurally stable [29, 31]. Morse-Smale flows and Anosov flows are Axiom A flows with the strong transversality condition. There are Axiom A flows other than Morse-Smale flows and Anosov flows that satisfy the strong transversality condition [7]. **Stability conjecture**: A C^r -flow is structurally stable (resp. Ω -stable) if and only if it is an Axiom A flow with the strong transversality condition (resp. the no cycle condition). S. Newhouse, V. A. Pliss, Robinson, R. Mañé, S. D. Liao,

and A. Sannami have made important contributions to the study of the stability conjecture.

Neither the set of all Axiom A flows nor the set of all Ω -stable flows nor the set of structurally stable flows is dense in $\mathcal{X}^r(M)$ if $\dim M > 2$ (R. Abraham and Smale, Newhouse). However, the set of all structurally stable flows is dense in $\mathcal{X}^r(M)$ in the C^0 topology (M. Shub, M. M. C. de Oliverira).

(4) A diffeomorphism $f \in \text{Diff}^r(M)$ is called a **Morse-Smale diffeomorphism** if the following conditions are satisfied: (i) The nonwandering set Ω is a finite set, and hence it consists of periodic points; (ii) all periodic points are hyperbolic; (iii) for each pair $p, q \in \Omega$, $W^s(p)$ intersects $W^u(q)$ transversally. The Morse-Smale diffeomorphisms on the unit circle S^1 are exactly the structurally stable ones described in Section I(1). A Morse-Smale diffeomorphism is structurally stable [25]. The set of all Morse-Smale diffeomorphisms on M is open but not dense in $\text{Diff}^r(M)$ if $\dim M \geq 2$ (Palis, Smale). However, it contains a dense open subset of the set of all time-one mappings of the flows generated by gradient vector fields. In particular, any closed manifold admits Morse-Smale diffeomorphisms and hence structurally stable diffeomorphisms. For a Morse-Smale diffeomorphism, the Morse inequalities hold [24].

Let Λ be a closed invariant set of $f \in \text{Diff}^r(M)$. Λ is called **hyperbolic** if the following conditions are satisfied: (i) There is a splitting $T_x(M) = E_x^s \oplus E_x^u$ of the tangent space $T_x(M)$ for each $x \in \Lambda$, which depends continuously on $x \in \Lambda$; (ii) $df_x(E_x^s) = E_{f(x)}^s$ and $df_x(E_x^u) = E_{f(x)}^u$ for all $x \in \Lambda$; (iii) There is a Riemannian metric on M and constants $c > 0$, $0 < \lambda < 1$ such that, for any integer $m > 0$ and $x \in \Lambda$, $\|df_x^m(w)\| \leq c\lambda^m\|w\|$ when $w \in E_x^s$ and $\|df_x^{-m}(w)\| \leq c\lambda^m\|w\|$ when $w \in E_x^u$. A diffeomorphism $f \in \text{Diff}^r(M)$ is called an **Anosov diffeomorphism** if M itself is hyperbolic for f . Manifolds which admit Anosov diffeomorphisms are restricted (Hirsch, K. Shiraiwa). Examples of Anosov diffeomorphisms are given as follows: Let $L: \mathbf{R}^n \rightarrow \mathbf{R}^n$ be a hyperbolic linear automorphism with $L(\mathbf{Z}^n) = \mathbf{Z}^n$, where \mathbf{Z}^n is the discrete subgroup of \mathbf{R}^n consisting of all elements with integral coordinates. Then L induces an automorphism $f: T^n \rightarrow T^n$ of the n -dimensional torus $T^n = \mathbf{R}^n/\mathbf{Z}^n$, which is an Anosov diffeomorphism. There are similar constructions using hyperbolic automorphisms of simply connected nilpotent Lie groups and their uniform discrete subgroups (Smale). A homeomorphism $h: X \rightarrow X$ of a metric space X is called **expansive** if there is a constant $\varepsilon > 0$ such that $x, y \in X$ and $d(h^n(x), h^n(y)) < \varepsilon$ for all $n \in \mathbf{Z}$ imply $x = y$. An Anosov diffeomorphism is expansive. For Anosov diffeomorphisms,

theorems similar to those stated in Section J(2) hold. Franks, Newhouse, A. Manning, and J. N. Mather obtained important results concerning Anosov diffeomorphisms.

A diffeomorphism $f \in \text{Diff}^r(M)$ is called an **Axiom A diffeomorphism** if the following conditions are satisfied: A(a) The nonwandering set Ω is hyperbolic; A(b) the set of all periodic points is dense in Ω . There are examples that satisfy Axiom A(a) but not Axiom A(b) (A. Dankner, M. Kurata). For $x \in M$, put $W^s(x) = \{y \in M \mid d(f^n(x), f^n(y)) \rightarrow 0 \text{ as } n \rightarrow \infty\}$ and $W^u(x) = \{y \in M \mid d(f^{-n}(x), f^{-n}(y)) \rightarrow 0 \text{ as } n \rightarrow \infty\}$. We call $W^s(x)$ (resp. $W^u(x)$) the **stable** (resp. **unstable**) manifold of f at x . For a subset A of M , put $W^\sigma(A) = \bigcup_{x \in A} W^\sigma(x)$ ($\sigma = s, u$). If f is an Axiom A diffeomorphism, $W^s(x)$ and $W^u(x)$ are injectively immersed submanifolds of M . If x is a hyperbolic periodic point, $W^s(x)$ and $W^u(x)$ defined here coincide with those defined before. For Axiom A diffeomorphisms, notions such as spectral decomposition, basic sets, the strong transversality condition, and the no cycle condition are defined similarly, and theorems similar to those stated in Section J(3) hold.

Let $h: X \rightarrow X$ be a homeomorphism of a metric space X and $\alpha, \beta > 0$. A sequence $\{x_i\}_{i \in \mathbf{Z}}$ of points in X is an **α -pseudo-orbit** if $d(h(x_i), x_{i+1}) < \alpha$ for all $i \in \mathbf{Z}$. We say that $\{x_i\}_{i \in \mathbf{Z}}$ is **β -shadowed** (or **β -traced**) by a point $x \in X$ if $d(h^i(x), x_i) \leq \beta$ for all $i \in \mathbf{Z}$. We say that h has the **pseudo-orbit tracing property** if for any $\beta > 0$ there exists an $\alpha > 0$ so that any α -pseudo-orbit is β -shadowed by some point. If f is an Axiom A diffeomorphism, then $f|_{\Omega(f)}: \Omega(f) \rightarrow \Omega(f)$ has the pseudo-orbit tracing property (Bowen). Axiom A diffeomorphisms with the strong transversality condition (in particular, Anosov diffeomorphisms) have the pseudo-orbit tracing property (Bowen, K. Sawada, K. Kato, A. Morimoto).

(5) Let S be a discrete topological space with n elements ($n \geq 2$) and $\Sigma(S) = \prod_{i \in \mathbf{Z}} S_i$ the doubly infinite product of copies S_i of S with the product topology. A point $x = (x_i)_{i \in \mathbf{Z}}$ of $\Sigma(S)$ is a doubly infinite sequence of points in S . $\Sigma(S)$ is a totally disconnected, perfect, compact, metrizable space (i.e., a Cantor set). Let $\sigma: \Sigma(S) \rightarrow \Sigma(S)$ be a mapping defined by $\sigma(x) = y$, $x = (x_i)$, $y = (y_i)$, $y_i = x_{i+1}$ ($i \in \mathbf{Z}$). Then σ is a homeomorphism of $\Sigma(S)$ and is called the **shift automorphism** with n symbols.

Define a diffeomorphism $f: S^2 \rightarrow S^2$ of the 2-dimensional sphere S^2 as follows. We consider S^2 as $\mathbf{R}^2 \cup \{\infty\}$. Let U be a region of \mathbf{R}^2 consisting of the rectangle $R (= ABCD)$ and an upper and a lower cap as shown in Fig. 1. f maps U into itself as in Fig. 1 ($f(A) = A'$, $f(B) = B'$, $f(C) = C'$, $f(D) = D'$, and so on). Here, $R \cap f(R) = P \cup Q$ is the union of two rectangles,

$P'=f^{-1}(P)$ and $Q'=f^{-1}(Q)$ are rectangles, and f is affine on each of P' and Q' (stretching in the vertical direction and contracting in the horizontal direction). The lower cap is contracted into itself, and every point in it tends to a sink x_0 . The upper cap is mapped into the lower one and stays there. The point ∞ is the only source, and $\lim_{n\rightarrow\infty} f^{-n}(x)=\infty$ for all $x\in\mathbb{R}^2-U$. Thus the nonwandering set of f consists of a sink x_0 , a source ∞ , and $\Lambda=\bigcap_{n\in\mathbb{Z}} f^n(R)$. The mapping f constructed here is called the **horseshoe diffeomorphism**. It is an Axiom A diffeomorphism with the strong transversality condition (and hence structurally stable). The restriction $f|_\Lambda$ of f on a basic set Λ is topologically conjugate to the shift automorphism with two symbols. It is neither a Morse-Smale diffeomorphism nor an Anosov diffeomorphism [7].

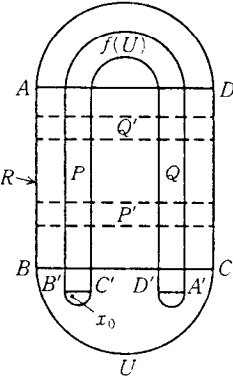


Fig. 1

A restriction of a shift automorphism on a closed invariant set is called a **subshift**. Let $A=(A_{ij})$ be an $n\times n$ matrix with (i,j) -component $A_{ij}=0$ or 1 for all $i,j=1,\dots,n$. It is **irreducible** if for each i,j there is a positive integer m such that A^m has a nonzero (i,j) -component. Assume that A is irreducible and $S=\{1,\dots,n\}$. Put $\Sigma_A=\{x=(x_i)\in\Sigma(S)|A_{x_i,x_{i+1}}=1\text{ for all }i\in\mathbb{Z}\}$. Then Σ_A is a closed invariant set of $\Sigma(S)$. The restriction $\sigma_A=\sigma|_{\Sigma_A}:\Sigma_A\rightarrow\Sigma_A$ is called a **subshift of finite type** or a **Markov subshift**, and A is called the **transition matrix**. A topological classification of the subshifts of finite type has been investigated by Williams.

Let $f\in\text{Diff}^r(M)$ be an Axiom A diffeomorphism and Λ its basic set. Bowen constructed a \dagger Markov partition for $f|_\Lambda$ by generalizing the Sinai construction for Anosov diffeomorphisms. The Markov partition connects $f|_\Lambda$ with a suitable subshift of finite type and is applied to the study of Axiom A diffeomorphisms, especially to the ergodic theory of Axiom A diffeomorphisms (Bowen, Ruelle). A similar theory for flows was developed by Bowen, Ruelle, and M. E. Ratner.

(6) Let $x\in M$ be a hyperbolic fixed point of

$f\in\text{Diff}^r(M)$. A point $p\in W^s(x)\cap W^u(x)$ ($p\neq x$) is called a **homoclinic point**. If $W^s(x)$ and $W^u(x)$ intersect transversally at a homoclinic point p , then p is called a **transversal homoclinic point**. In a neighborhood of a transversal homoclinic point, there is a closed invariant set Λ of f^m for some positive integer m such that $f^m|_\Lambda$ is topologically conjugate to the shift automorphism with two symbols (Smale). There are generalizations of this theorem for semiflows by F. R. Marotto, Shiraiwa, and Kurata.

K. Topological Entropy and Zeta Functions

(1) The notion of topological entropy was first defined by R. L. Adler, A. G. Konheim, and M. H. McAndrew as an analog to measure-theoretic entropy. Let X be a compact topological space and α an open covering of X . Let $N(\alpha)$ be the minimum number of members of a subcovering of α . Let $f:X\rightarrow X$ be a continuous mapping and α an open covering of X . Then $\lim_{n\rightarrow\infty} (1/n)\log N(\alpha\vee f^{-1}\alpha\vee\dots\vee f^{n-1}\alpha)$ exists, where $\alpha\vee f^{-1}\alpha\vee\dots\vee f^{n-1}\alpha$ is the open covering $\{A_0\cap f^{-1}(A_1)\cap\dots\cap f^{n-1}(A_{n-1})|A_0,A_1,\dots,A_{n-1}\in\alpha\}$. We denote the above limit by $h(f,\alpha)$ and call it the **topological entropy of f with respect to α** . The **topological entropy $h(f)$** of f is defined as the $\sup h(f,\alpha)$, where the supremum is taken over all open coverings α of X . Now assume that X is a compact metric space. For an open covering α of X , put $\text{diam}\alpha=\sup\{\text{diam} A|A\in\alpha\}$, where $\text{diam} A$ is the diameter of A . If $\{\alpha_n\}_{n\geq 1}$ is a sequence of open coverings of X such that $\text{diam}\alpha_n\rightarrow 0$ as $n\rightarrow\infty$, then $h(f,\alpha_n)\rightarrow h(f)$ as $n\rightarrow\infty$. The topological entropy of the shift automorphism with n symbols is equal to $\log n$. Let $L:\mathbb{R}^n\rightarrow\mathbb{R}^n$ be a linear mapping with $L(\mathbb{Z}^n)\subset L(\mathbb{Z}^n)$ and $\lambda_1,\dots,\lambda_n$ its eigenvalues. Let $f:T^n\rightarrow T^n$ be the induced endomorphism of the n -dimensional torus $T^n=\mathbb{R}^n/\mathbb{Z}^n$. Then the topological entropy is given by $h(f)=\sum_{|\lambda_i|>1}\log|\lambda_i|$.

Let $f:X\rightarrow X$ be a continuous mapping and $M(f)$ the set of all $\dagger f$ -invariant probability measures on the Borel sets of X . For each $\mu\in M(f)$, denote by $h_\mu(f)$ the (measure-theoretic) entropy of f with respect to μ . Then $h(f)=\sup\{h_\mu(f)|\mu\in M(f)\}$ (E. I. Dinaburg, T. N. T. Goodman, L. W. Goodwyn). The following properties hold: (i) If $f:X\rightarrow X$ is a homeomorphism, then $h(f^n)=|n|h(f)$ for all $n\in\mathbb{Z}$. (ii) If (X,φ) is a continuous flow, then $h(\varphi_t)=|t|h(\varphi_1)$ for all $t\in\mathbb{R}$. (iii) Let $f_i:X_i\rightarrow X_i$ ($i=1,2$) be continuous. Then $h(f_1\times f_2)=h(f_1)+h(f_2)$. (iv) Let $f_i:X_i\rightarrow X_i$ ($i=1,2$) be continuous. If there is a continuous mapping $g:X_1\rightarrow X_2$ such that $g(X_1)=X_2$ and $g\circ f_1=f_2\circ g$, then $h(f_1)\geq h(f_2)$. In particular, if f_1 and f_2 are topologically conjugate, then $h(f_1)=h(f_2)$.

$h(f_2)$. (v) If $f: X \rightarrow X$ is a homeomorphism and Ω is the nonwandering set of f , then $h(f) = h(f|_{\Omega})$. In particular, Ω -conjugate homeomorphisms have the same topological entropy (Bowen). (vi) If $f: M \rightarrow M$ is an Axiom A diffeomorphism of a closed C^∞ -manifold M , then $h(f) = \limsup_{n \rightarrow \infty} (1/n) \log N_n(f)$, where $N_n(f)$ is the number of all periodic points of period n . And $h(f) = 0$ if and only if the nonwandering set is finite (Bowen).

Bowen gave alternative definitions of the topological entropy and generalized it for uniformly continuous mappings of (not necessarily compact) metric spaces. If $f: M \rightarrow M$ is a C^1 -mapping of an n -dimensional Riemannian manifold M , then $h(f) \leq \max\{0, n \log \sup\{\|df_x\| \mid x \in M\}\}$, where $\|df_x\|$ is the norm of $df_x: T_x(M) \rightarrow T_{f(x)}(M)$, and hence $h(f)$ is finite (Bowen, S. Ito).

(2) Let $f: M \rightarrow M$ be a continuous mapping of a closed C^∞ -manifold M and $f_*: H_*(M) \rightarrow H_*(M)$ (resp. $f_{*1}: H_1(M) \rightarrow H_1(M)$) the \dagger induced homomorphism if f on the \dagger homology group $H_*(M)$ (resp. the \dagger first homology group $H_1(M)$) with coefficients in \mathbf{R} . The **spectral radius** $s(L)$ of a linear mapping $L: E \rightarrow E$ of a real vector space E is the maximum of the absolute values of the eigenvalues of L . The following **entropy conjecture** is still open: If $f: M \rightarrow M$ is a C^1 -diffeomorphism (C^1 -mapping), then $h(f) \geq \log s(f_*)$. Concerning the entropy conjecture, the following are known: (i) The conjecture holds for a dense open set of $\text{Diff}^r(M)$ in the C^0 -topology (Shub); (ii) for any continuous mapping f , $h(f) \geq \log s(f_{*1})$ (Manning); (iii) the conjecture fails for a homeomorphism (Pugh); (iv) the conjecture holds for an Axiom A diffeomorphism with the no cycle condition (Shub and Williams); (v) the conjecture holds for any continuous mapping of the n -dimensional torus (M. Misiurewicz, F. Przytycki); (vi) for a C^1 -mapping $f: M \rightarrow M$, $h(f) \geq \log |\deg f|$, where $\deg f$ is the \dagger mapping degree of f (Misiurewicz, Przytycki).

(3) M. Artin and B. Mazur first defined the zeta function of a diffeomorphism by analogy to \dagger Weil's zeta function. Let $f: X \rightarrow X$ be a homeomorphism of a compact metric space X . Assume that the number $N_m = N_m(f)$ of all periodic points of period m is finite for all m . Put $\zeta(t) = \exp(\sum_{m=1}^{\infty} N_m t^m / m)$ and call it the **zeta function** of f . For any closed C^∞ -manifold M , there is a dense set E of $\text{Diff}^r(M)$ such that $\tilde{N}_m(f) \leq ck^m$ for $f \in E$, where $\tilde{N}_m(f)$ is the number of isolated periodic points of period m and c and k are positive constants depending only on f (Artin, Mazur). Hence, for such $f \in E$, the series $\exp(\sum_{m=1}^{\infty} \tilde{N}_m(f) t^m / m)$ has a positive radius of convergence. Originally, Artin and Mazur called this the zeta function of f .

Let $\sigma_A: \Sigma_A \rightarrow \Sigma_A$ be a subshift of finite type with transition matrix A ; then $\zeta(t) = 1/\det(E - tA)$, where E is the unit matrix (Bowen and O. E. Lanford III). The zeta function of an Axiom A diffeomorphism has a positive radius of convergence and is a rational function (K. Meyer, J. Guckenheimer, Manning).

Let (M, φ) be a nonsingular C^r -flow on a closed C^∞ -manifold M . Let Γ be the set of all closed orbits and $l(\gamma)$ the smallest positive period of $\gamma \in \Gamma$. Smale defined the **zeta function** of (M, φ) by $Z(s) = \prod_{\gamma \in \Gamma} \prod_{k=0}^{\infty} (1 - [\exp l(\gamma)]^{-s-k})$. If (M, φ) is the geodesic flow on a surface of constant negative curvature, $Z(s)$ reduces to the \dagger Selberg zeta function.

There are generalizations and modifications of the notion of zeta function by Ruelle and Franks.

L. Classical Dynamical Systems

(1) Let M be a C^∞ -manifold without boundary. A C^r -flow (M, φ) or a C^r -diffeomorphism $f: M \rightarrow M$ with a smooth invariant measure is called a **classical dynamical system**. Most important classical dynamical systems are Hamiltonian or Lagrangian systems. In the modern formulation, a **Hamiltonian system** consists of M , a **symplectic form** ω on M (i.e., a \dagger nondegenerate \dagger closed 2-form), and the vector field X_H on M defined by a C^{r+1} -function $H: M \rightarrow \mathbf{R}$. We call X_H the **Hamiltonian vector field** with **energy function** H . Let (M, ω, X_H) be a Hamiltonian system. Then the following hold: (i) M is of even dimension and there is a system of local coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$ such that $\omega = \sum_{i=1}^n dq^i \wedge dp_i$ (J. G. Darboux). In these coordinates X_H is expressed by \dagger Hamilton's equations $dq^i/dt = \partial H / \partial p_i$, $dp_i/dt = -\partial H / \partial q^i$, $i = 1, \dots, n$; (ii) the smooth measure defined by the \dagger volume element $\Omega = ((-1)^{n(n-1)/2}/n!) \omega^n$ is an invariant measure for the flow generated by X_H (J. Liouville); (iii) the energy function H is constant along any trajectory of X_H . Especially, $H^{-1}(e)$ ($e \in \mathbf{R}$) is an invariant set for each e and is called an **energy surface**. Energy surfaces are submanifolds of codimension one for almost all $e \in \mathbf{R}$.

Important examples of Hamiltonian systems are given as follows. Let Q be an n -dimensional manifold and $T^*(Q)$ the \dagger cotangent bundle of Q . Then $T^*(Q)$ has a canonical symplectic form ω_0 . We call Q a **configuration space** and $T^*(Q)$ a **momentum phase space**. For any C^{r+1} -function H on $T^*(Q)$, we have a Hamiltonian system $(T^*(Q), \omega_0, X_H)$ in which X_H is of class C^r . The Hamiltonian formalism is translated into Lagrangian formalism by using the \dagger tangent bundle $T(Q)$ instead of $T^*(Q)$. $T(Q)$ is called a **velocity phase space**.

For a given differentiable function L on $T(Q)$, the **energy function** E on $T(Q)$ and the **Lagrangian vector field** X_E on $T(Q)$ are constructed. If L satisfies a certain condition, there is a system of local coordinates $(q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n)$ such that X_E is expressed in these coordinates by the \dagger Euler-Lagrange equations $dq^i/dt = \dot{q}^i$, $d/dt(\partial L/\partial \dot{q}^i) = \partial L/\partial q^i$, $i = 1, \dots, n$. The projection of an integral curve (i.e., an orbit) of X_E into Q is called a base integral curve. Under a suitable condition, the base integral curve of a Lagrangian (or a Hamiltonian) system is the geodesic of the Jacobi metric on Q up to a reparametrization, and the restriction of the flow generated by X_X on an energy surface is the geodesic flow described below.

(2) Let M be a \dagger complete Riemannian manifold of class C^∞ and $S(M) = \{v \in T(M) \mid \|v\| = 1\}$. Let $\pi: S(M) \rightarrow M$ be the projection defined by $\pi(v) = x$ for $v \in T_x(M)$. Then $S(M)$ is a \dagger sphere bundle over M , which is called the **unit tangent sphere bundle** over M . Each $v \in S(M)$ determines a unique \dagger geodesic $C_v: \mathbf{R} \rightarrow M$ such that $C_v(0) = \pi(v)$ and the \dagger tangent vector $C'_v(0)$ to C_v at $t = 0$ is equal to v . Let v_t be the tangent vector $C'_v(t)$ to C_v at t for $t \in \mathbf{R}$. Then $\|v_t\| = 1$ and $\pi(v_t) = C_v(t)$ for all $t \in \mathbf{R}$. Define a mapping $\varphi: S(M) \times \mathbf{R} \rightarrow S(M)$ by $\varphi(v, t) = v_t$. Then $(S(M), \varphi)$ is a C^∞ -flow, which is called the **geodesic flow** on M . By the classical Liouville theorem, a geodesic flow has a smooth invariant measure.

(3) Let $T^n = \mathbf{R}^n/\mathbf{Z}^n$ be the n -dimensional torus and $\omega_1, \dots, \omega_n$ real constants. Define a mapping $\varphi: T^n \times \mathbf{R} \rightarrow T^n$ by $\varphi([x_1, \dots, x_n], t) = [x_1 + \omega_1 t, \dots, x_n + \omega_n t]$, where $[x_1, \dots, x_n] \in \mathbf{R}^n/\mathbf{Z}^n = T^n$ is the residue class of $(x_1, \dots, x_n) \in \mathbf{R}^n$ modulo \mathbf{Z}^n . Then (T^n, φ) is a C^∞ -flow. We call it the **translation flow with frequencies** $\omega_1, \dots, \omega_n$. If $\omega_1, \dots, \omega_n$ are linearly independent over \mathbf{Z} , they are called **independent**. Every orbit of the translation flow is dense if and only if its frequencies are independent (Poincaré, H. Weyl). A translational flow with independent frequencies is called **quasiperiodic**.

Let (M, ω, X_H) be a Hamiltonian system on a $2n$ -dimensional manifold M . Under a certain condition, there exist an open set U of M and a diffeomorphism $f: U \rightarrow T^n \times \mathbf{R}^n$ such that the following holds: Identify U with $T^n \times \mathbf{R}^n$ by f , and let $(q^1, \dots, q^n, p_1, \dots, p_n)$ be the coordinates of $T^n \times \mathbf{R}^n$. Then the energy function H is independent of $q = (q^1, \dots, q^n)$ so that Hamilton's equations becomes $dq^i/dt = \partial H/\partial p_i$, $dp_i/dt = -\partial H/\partial q^i = 0$ for $i = 1, \dots, n$. Therefore the solutions are given by $q^i(t) = (\partial H/\partial p_i(c))t + q^i(0)$ modulo \mathbf{Z} , $p_i(t) = p_i(0) = c_i$, $i = 1, \dots, n$, where $c = (c_1, \dots, c_n)$. Therefore $N_c = f^{-1}(T^n \times \{c\})$ is an **invariant torus** (i.e., an invariant set diffeomorphic to T^n) of (the flow generated by) X_H for all $c \in \mathbf{R}^n$, and the restriction flow on

N_c is the translational flow with frequencies $\omega_1 = \partial H/\partial p_1(c), \dots, \omega_n = \partial H/\partial p_n(c)$ (Arnold). Assume further that the Hessian $\det(\partial^2 H/\partial p_i \partial p_j)$ does not vanish at c and $\omega_1, \dots, \omega_n$ are independent. Let \tilde{H} be an energy function obtained by adding a sufficiently small perturbation to H . Then for almost all c' near c , there exists an invariant torus \tilde{N} of $X_{\tilde{H}}$ near N_c such that the restriction flow on \tilde{N} is differentially equivalent to the translational flow with the same frequencies (Kolmogorov, Arnold, Moser).

(4) Generic properties for Hamiltonian systems were investigated by M. Buchner, Markus, Meyer, Pugh, Robinson, Takens, and Newhouse.

M. Bifurcation

(1) Consider a differential equation with a parameter. For example, let X be a domain of \mathbf{R}^n , $J = (-1, 1)$, and $f: J \times X \rightarrow \mathbf{R}^n$ a C^r -mapping. For each $\mu \in J$, define $f_\mu: X \rightarrow \mathbf{R}^n$ by $f_\mu(x) = f(\mu, x)$, $x \in X$. Consider the differential equation

$$dx/dt = f_\mu(x), \quad x \in X \text{ and } \mu \in J. \tag{7}$$

As μ varies, the topological structure of the phase portrait of (7) may change. Suppose that there exists $\mu_0 \in J$ such that the topological structure of the phase portrait of (7) changes at $\mu = \mu_0$ but remains the same when $\mu_0 - \varepsilon < \mu < \mu_0$ or $\mu_0 < \mu < \mu_0 + \varepsilon$ for some $\varepsilon > 0$. Then μ_0 is called a **bifurcation point** of (7).

Hopf bifurcation: Assume that $X = \mathbf{R}^2$ and the origin $0 \in \mathbf{R}^2$ is a singular point of (7) for all $\mu \in J$. Assume further that the Jacobian matrix of f_μ at $0 \in \mathbf{R}^2$ has two distinct complex conjugate eigenvalues $\lambda(\mu)$ and $\bar{\lambda}(\mu)$ such that the real part $\text{Re } \lambda(\mu)$ of $\lambda(\mu)$ is positive when $\mu > 0$, zero when $\mu = 0$, and negative when $\mu < 0$. Then $0 \in \mathbf{R}^2$ is a sink for $\mu < 0$ and a source for $\mu > 0$. Now assume further that $d/d\mu(\text{Re } \lambda(\mu))|_{\mu=0}$ is positive and $0 \in \mathbf{R}^2$ is a "vague attractor." Then $0 \in J$ is a bifurcation point, and there exists an asymptotically stable closed orbit for (7) near and around $0 \in \mathbf{R}^2$ which depends continuously on μ for $\mu > 0$ [37]. Thus a sink of (7) ($\mu < 0$) changes to a source and an asymptotically stable closed orbit ($\mu > 0$) when μ changes its sign. The Hopf bifurcation theorem can be generalized to a higher-dimensional case, and there is a diffeomorphism version of the theorem.

(2) Bifurcations in more general settings have been investigated by many mathematicians, including Thom, Arnold, R. J. Sacker, G. R. Sell, D. H. Sattinger, G. Iooss, Ruelle, and Takens. Generic bifurcations of dynamical systems have been investigated by J. Soto-

mayer, Meyer, P. Brunovský, and others; bifurcations of Morse-Smale systems by Newhouse, Palis, Peixoto, and S. Matsumoto; and bifurcations of Axiom A diffeomorphisms by Newhouse and Palis.

N. Miscellaneous Topics

(1) Let S^3 be the 3-dimensional sphere and $V \in \mathcal{X}^1(S^3)$. H. Seifert proved that if V is sufficiently close (in C^0 topology) to a nonsingular vector field tangent to the fibers of the 1 Hopf fibration $S^3 \rightarrow S^2$, then V has a closed orbit. He conjectured that every nonsingular vector field $V \in \mathcal{X}^1(S^3)$ had a closed orbit. (**Seifert conjecture**; \rightarrow 154 Foliations D). If a nonsingular vector field $V \in \mathcal{X}^1(S^3)$ is transverse to a codimension 1 foliation of class C^2 , then it has a closed orbit (S. P. Novikov). Let M be a 3-dimensional C^∞ -manifold. Then there exists a nonsingular vector field $V \in \mathcal{X}^1(M)$ with no closed orbit in any 1 homotopy class of a nonsingular vector field on M (P. A. Schweitzer). Thus the Seifert conjecture fails for a vector field of class C^1 , but the conjecture for a vector field of class C^r ($r \geq 2$) is an open problem. Related work has been done by Fuller, H. Chu, and A. Weinstein.

(2) Let M be a closed connected C^∞ -manifold. A C^r -flow (M, φ) (resp. $f \in \text{Diff}^r(M)$) is a **minimal flow** (resp. a **minimal diffeomorphism**) if M itself is a minimal set. If M admits a 1 locally free S^1 -action of class C^∞ , then it admits a minimal C^∞ -diffeomorphism, and if M admits a locally free special (in particular, a 1 free) T^2 -action of class C^∞ , then it admits a minimal C^∞ -flow (A. Fathi, Herman, A. B. Katok). Open problems: What are the topological properties of the manifolds admitting minimal flows? Does S^3 admit a minimal flow?

(3) E. N. Lorenz studied numerical solutions of the following nonlinear equations in \mathbf{R}^3 which arose from the convection equation: $dx/dt = -\sigma x + \sigma y$, $dy/dt = -xz + rx - y$, $dz/dt = xy/bz$. When $\sigma = 10$, $r = 28$, and $b = 8/3$, he found irregular behavior in this dynamical system. R. M. May studied numerical solutions of the following difference equation in connection with the growth of biological populations with nonoverlapping generations: $x_{n+1} = ax_n(1 - x_n)$, $x_n \in [0, 1]$ ($1 < a < 4$). He found that the dynamical structure of the above difference equation was delicate and complicated. Y. Ueda and H. Kawakami also found similar phenomena in their numerical study of Duffing's equation of the type $d^2x/dt^2 + k dx/dt + x^3 = B \cos t$. The phenomena observed in these investigations were called **chaos**, which exhibits **strange attractors**.

These investigations have attracted the

attention of many mathematicians and scientists. For 1-dimensional semidynamical systems such as May's equation, T. Y. Li, J. A. Yorke, A. N. Sharkovskii, J. W. Milnor, Thurston, and many others have obtained notable results, while for Lorenz's equation we have results by Ruelle, Guckenheimer, Williams, Sinai, and many others. Chaos arising from discretization of differential equations has been studied by M. Yamaguti, S. Ushiki, and others (\rightarrow 433 Turbulence and Chaos).

References

- [1] H. Poincaré, Mémoire sur les courbes définies par une équation différentielle, J. Math. Pures Appl., (3) 7 (1881), 375–422; 8 (1882), 251–296. (Oeuvres I, Gauthier-Villars, 1928, 3–84.)
- [2] H. Poincaré, Sur les courbes définies par les équations différentielles, J. Math. Pures Appl., (4) 1 (1885), 167–244; 2 (1886), 151–217. (Oeuvres I, 90–158, 167–222.)
- [3] H. Poincaré, Les méthodes nouvelles de la mécanique céleste. I, Solutions périodiques, Non-existence des intégrals uniformes, Solutions asymptotiques. II, Méthodes de MM. Newcomb, Gylden, Lindstedt et Bohlin. III, Invariants intégraux, Solutions périodiques du deuxième genre, Solutions doublement asymptotiques. Gauthier-Villars, I, 1892; II, 1893; III, 1899; English translation, New methods of celestial mechanics I–III, Clearinghouse for Federal Scientific and Technical Information, Springfield, 1967.
- [4] G. D. Birkhoff, Dynamical systems, Amer. Math. Soc. Colloq. Publ., 1927.
- [5] A. A. Andronov and L. S. Pontryagin, Systèmes grossiers, Dokl. Akad. Nauk SSSR, 14 (1937), 247–250.
- [6] S. Lefschetz, Differential equations: geometric theory, Interscience, 1957.
- [7] S. Smale, Differentiable dynamical systems, Bull. Amer. Math. Soc., 73 (1967), 747–817.
- [8] D. V. Anosov, Geodesic flows on closed Riemannian manifolds with negative curvature, Proc. Steklov Inst. Math., 90 (1967), 1–235.
- [9] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms, Lecture notes in math. 470, Springer, 1975.
- [10] D. Ruelle and F. Takens, On the nature of turbulence, Comm. Math. Phys., 20 (1971), 167–192.
- [11] J. Moser, Stable and random motions in dynamical systems, Ann. Math. Studies 77, Princeton Univ. Press, 1973.
- [12] L. H. Loomis and S. Sternberg, Advanced calculus, Addison-Wesley, 1968.

- [13] V. V. Nemytskii and V. V. Stepanov, Qualitative theory of differential equations, Princeton Univ. Press, 1960. (Original in Russian, 1947.)
- [14] N. P. Bhatia and G. P. Szegő, Stability theory of dynamical systems, Springer, 1970.
- [15] W. H. Gottschalk and G. A. Hedlund, Topological dynamics, Amer. Math. Soc. Colloq. Publ., 1955.
- [16] R. Abraham and J. W. Robbin, Transversal mappings and flows, Benjamin, 1967.
- [17] L. Markus, Lectures in differentiable dynamics, revised edition, Regional Conf. Series in Math. 3, Amer. Math. Soc., 1980.
- [18] Z. Nitecki, Differentiable dynamics, MIT Press, 1971.
- [19] M. Shub, Stabilité globale des systèmes dynamiques, Astérisque, Soc. Math. France, 56 (1978).
- [20] M. C. Irwin, Smooth dynamical systems, Academic Press, 1980.
- [21] S. Smale, Stable manifolds for differential equations and diffeomorphisms, Ann. Scuola Norm. Sup. Pisa, (3) 17 (1963), 97–116.
- [22] I. Kupka, Contribution à la théorie des champs génériques, Contributions to Differential Equations, 2 (1963), 457–482.
- [23] C. Pugh, An improved closing lemma and a general density theorem, Amer. J. Math., 89 (1967), 1010–1021.
- [24] S. Smale, Morse inequalities for a dynamical system, Bull. Amer. Math. Soc., 66 (1960), 43–49.
- [25] J. Palis and S. Smale, Structural stability theorems, Amer. Math. Soc. Proc. Symp. Pure Math., 14 (1970), 223–231.
- [26] M. W. Hirsch and C. C. Pugh, Stable manifolds and hyperbolic sets, Amer. Math. Soc. Proc. Symp. Pure Math., 14 (1970), 133–163.
- [27] S. Smale, The Ω -stability theorem, Amer. Math. Soc. Proc. Symp. Pure Math., 14 (1970), 289–297.
- [28] S. Smale, Notes on differentiable dynamical systems, Amer. Math. Soc. Proc. Symp. Pure Math., 14 (1970), 277–287.
- [29] J. W. Robbin, A structural stability theorem, Ann. Math., (2) 94 (1971), 447–493.
- [30] J. W. Robbin, Topological conjugacy and structural stability for discrete dynamical systems, Bull. Amer. Math. Soc., 78 (1972), 923–952.
- [31] R. C. Robinson, Structural stability of C^1 flows, Lecture notes in math. 468, Springer, 1975, 262–277.
- [32] M. Shub, Dynamical systems, filtrations and entropy, Bull. Amer. Math. Soc., 80 (1974), 27–41.
- [33] M. W. Hirsch, C. C. Pugh, and M. Shub, Invariant manifolds, Lecture notes in math. 583, Springer, 1977.

- [34] R. Bowen, On Axiom A diffeomorphisms, Regional Conf. Series in Math. 35, Amer. Math. Soc., 1978.
- [35] R. Abraham and J. E. Marsden, Foundations of mechanics, second edition, Benjamin, 1978.
- [36] V. I. Arnold and A. Avez, Problèmes ergodiques de la mécanique classique, Gauthier-Villars, 1967; English translation, Ergodic problems of classical mechanics, Benjamin, 1968.
- [37] J. E. Marsden and M. McCracken, The Hopf bifurcation and its applications, Applied math. sci. 19, Springer, 1976.

127 (XIX.8) Dynamic Programming

A. General Remarks

There are two types of multistage [†]decision processes. In one of them, an outcome of the whole process is determined at the final stage without any consideration of the outcome for each intermediate stage. The [†]extensive form of a [†]game is of this type. In the other type, an outcome is assigned at each stage of a multistage decision process. The theory of **dynamic programming**, dealing with this latter type, has been developed by R. Bellman and others since 1950 and is now one of the fundamental branches of mathematical programming, along with the theories of linear and nonlinear programming. The following examples illustrate some features of multistage decision processes.

Multistage allocation process. We are given a quantity $x > 0$ that can be divided into two parts y and $(x - y)$. From y we obtain a return $g(y)$, and from $(x - y)$ a return $h(x - y)$. In so doing, we expend a certain amount of our original resources and are left with a new quantity, $ay + b(x - y)$, $0 < a, b < 1$, with which the process is continued. How do we proceed so as to maximize the total return obtained in a finite or unbounded number of stages?

Multistage choice process. Suppose that we possess two gold mines A and B , the first of which contains an amount x of gold, while the second contains an amount y . In addition, we have a single gold mining machine with the property that if used to mine gold in the mine A , there is a probability P_1 that it will mine a fraction r_1 of the gold there and remain in working order, and a probability $1 - P_1$ that it will mine no gold and be damaged beyond repair. Similarly, the mine B has associated

with it the corresponding probabilities P_2 and $1 - P_2$ and fraction r_2 . How do we proceed in order to maximize the total amount of gold before the machine is defunct?

These two processes have the following features in common: (1) In each case we have a physical system characterized in any state by a small set of parameters, the **state variables**. (2) In each state of either process we have a choice of a number of **decisions**. (3) The effect of a decision is a transformation of the state variables. (4) The past history of the system is of no importance in determining future actions (*Markov property). (5) The purpose of the process is to maximize some function of the state variables.

A **policy** is a rule for making decisions that yields an allowable sequence of decisions; an **optimal policy** is a policy that maximizes a preassigned function of the final state variables. A convenient term for this preassigned function of the final state variables is **criterion function**. One of the characteristic features of Bellman's methodology of dynamic programming is the appeal to the **principle of optimality**: An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

In the multistage allocation process the state variables are x (the quantity of resources) and z (the return obtained up to the current stage). The decision at any stage consists of an allocation of a quantity $0 \leq y \leq x$. This decision has the effect of transforming x into $ay + b(x - y)$ and z into $z + g(y) + h(x - y)$. The purpose of the process is to maximize the final value of z . Denote by $f_N(x)$ the N -stage return obtained starting from an initial state x and using an optimal policy. Then we have

$$f_1(x) = \max_{0 \leq y \leq x} [g(y) + h(x - y)],$$

$$f_n(x) = \max_{0 \leq y \leq x} [g(y) + h(x - y) + f_{n-1}(ay + b(x - y))], \quad n \geq 2.$$

This recurrence relation yields a method for obtaining the sequence $\{f_n(x)\}$ inductively.

In the stochastic gold-mining process, the state variables are x and y (the present level of the two mines) and z (the amount of gold mined to date). The decision at any stage consists of a choice of A and B . If A is chosen, (x, y) goes into $((1 - r_1)x, y)$ and z into $z + r_1x$, and if B is chosen, (x, y) goes into $(x, (1 - r_2)y)$ and z into $z + r_2y$. The purpose of the process is to maximize the expected value of z obtained before the machine becomes defunct.

Denote by $f(x, y)$ the expected amount of gold obtained using an optimal sequence of

choice. Then we have

$$f(x, y) = \max \left[\begin{array}{l} p_1 \{r_1x + f((1 - r_1)x, y)\} \\ p_2 \{r_2y + f(x, (1 - r_2)y)\} \end{array} \right].$$

The optimal policy can be described in the following way. We choose A or B according as $p_1r_1x/(1 - p_1)$ is greater or less than $p_2r_2y/(1 - p_2)$. We can choose either A or B if equality holds. After an operation according to such a choice, the machine may become defunct and terminate the process. If the machine is usable, then we can apply our policy to a new combination of the amounts of gold in A and B .

B. Discrete Deterministic Processes

By a deterministic process we mean a process in which the outcome of a decision is uniquely determined by the decision. We assume that the state of the system, apart from time dependence, is described in any stage by an M -dimensional vector p constrained to lie within some region D . Let $T = \{T_q\}$ (where q runs over a set S) be a set of transformations with the property that $p \in D$ implies that $T_q(p) \in D$ for all $q \in S$, i.e., any transformation T_q carries D into itself. The term "discrete" signifies here that we have a process consisting of a finite or denumerably infinite number of stages. A policy, for the finite process which we consider first, consists of a selection of N transformations in order, $P = (T_1, T_2, \dots, T_N)$, yielding successively the sequence of states $p_i = T_i(p_{i-1})$ ($i = 2, 3, \dots, N$) with $p_1 = T_1(p)$. These transformations are to be chosen to maximize a given function R of the final state p_N . Observe that the maximum value of $R(p_N)$, as determined by an optimal policy, will be a function of the initial vector p and the number N of stages only. Let us then define our basic auxiliary functions $f_N(p) = \max R(p_N) =$ the N -stage return obtained starting from an initial state p and using an optimal policy. This sequence is defined for $N = 1, 2, \dots$, and $p \in D$. The essential uses of the principle of optimality can be observed from the following two features. The first is the use of the **embedding principle**. The original process is embedded in a family of similar processes. In place of attempting to determine the characteristics of an optimal policy for an isolated process, we attempt to deduce the common properties of the set of optimal policies possessed by the members of the family. The second feature is the derivation of recurrence relations by which the functional equations connecting the members of the sequence $\{f_k(p)\}$ are established. Assume that we choose some transformation T_q as a result of our first decision, obtaining in this way a new state vector $T_q(p)$. The **maximum return**

from the following $k-1$ stages is, by definition, $f_{k-1}(T_q(p))$. It follows that, if we wish to maximize the total k -stage return, q must now be chosen to maximize this $(k-1)$ -stage return. The result is the basic recurrence relation $f_k(p) = \max_{q \in S} f_{k-1}(T_q(p))$, for $k \geq 2$, with $f_1(p) = \max_{q \in S} R(T_q(p))$. For the case of an unbounded process, the sequence $\{f_k(p)\}$ is replaced by a single function $f(p)$, the total return obtained by using an optimal policy starting from state p , and the recurrence relation is replaced by the functional equation $f(p) = \max_q f(T_q(p))$.

C. Discrete Stochastic Processes

We again consider a discrete process, but one in which the transformations are stochastic rather than deterministic. The initial vector p is transformed into a stochastic vector z with an associated distribution function $dG_q(p, z)$ dependent on p and the choice of q . We assume that z is known after the decision has been made and before the next decision is to be made. We agree to measure the value of a policy in terms of some average value of the function of the final state. Let us call this expected value the **return**. Beginning with the case of a finite process, we define $f_k(p)$ as before. The expected return as a result of the initial choice of T_q is therefore

$$\int_{z \in D} f_{k-1}(z) dG_q(p, z).$$

Consequently, the recurrence relation for the sequence $\{f_k(p)\}$ is

$$f_k(p) = \max_{q \in S} \int_{z \in D} f_{k-1}(z) dG_q(p, z), \quad k \geq 2,$$

with $f_1(p) = \max_{q \in S} \int_{z \in D} R(z) dG_q(p, z)$. Considering the unbounded process, we obtain the functional relation

$$f(p) = \max_{q \in S} \int_{z \in D} f(z) dG_q(p, z).$$

D. Continuous Deterministic Processes

A number of interesting processes require that decisions be made at each point of a continuum, such as a time interval. The simplest examples of processes of this character are provided by the calculus of variations. Let us denote by $f(p; T)$ the return obtained over a time interval $[0, T]$ starting from the initial state p and employing an optimal policy. Although we consider the process as one consisting of choices made at each point t on $[0, T]$, it is better to begin with the concept of choosing policies (functions) over intervals, and then pass to the limit as these intervals shrink to points. Let d be an allowable deci-

sion made over the interval $[0, S]$, and let p_d be the state at S starting from the initial state p and employing d . The application of the principle of optimality suggests that

$$f(p; S+T) = \sup_D f(p_d, T), \quad (1)$$

where the supremum is taken over the set D of all allowable decisions d .

The limiting form of (1) as $S \rightarrow 0$ is a nonlinear partial differential equation (Bellman partial differential equation). This expression is important for use in actual analysis. For numerical purposes, S is kept nonzero but small. R. Bellman showed that it is possible to avoid many of the quite difficult rigorous details involved in this limiting procedure if we are interested only in the computational solution of variational processes.

E. Markovian Decision Processes

Consider a physical system which at any of the times $t=0, \Delta, 2\Delta, \dots$ must lie in one of the states S_1, S_2, \dots, S_N . Let $y_i(n)$ be the probability that the system is in S_i at times $n\Delta$, and let P_{ij} be the probability that the system is in state S_j at $t+\Delta$ if it is in state S_i at time t . We suppose that the transition probabilities P_{ij} are independent of t . We assume that the P_{ij} depend on a parameter q , which may be a vector, and that at each stage of the process q is to be chosen so as to maximize the probability that the system is in the state S_1 . We obtain the nonlinear system

$$\begin{aligned} y_1(n+1) &= \max_q \sum_{j=1}^N p_{1j}(q) y_j(n) \\ &= \sum_{i=1}^N y_i(n) p_{i1}(q^*); \end{aligned}$$

$$y_i(n+1) = \sum_{j=1}^N p_{ij}(q^*) y_j(n), \quad i=2, 3, \dots, N,$$

where $q^* = q^*(n)$ in the remaining $N-1$ equations is one of the values of q that maximize $y_1(n+1)$. There are similar processes that can be considered as continuous analogs of this type of decision process. These are called **Markovian decision processes** and were discussed by Bellman. There is, however, another type of Markovian decision process in which a reward is given at each stage. For each state S_i of the system there are k alternatives $1, 2, 3, \dots, k$. If we choose the alternative h among these k alternatives, then the transition probabilities p_{ij}^h ($j=1, 2, \dots, n$) are determined, and a reward r_{ij}^h is associated with each state S_j . Let us denote by $v_i(n)$ the total expected return obtained at the n th stage by appealing to an optimal policy when the initial state is S_i . Then the principle of optimality in the theory

of dynamic programming yields

$$v_i(n+1) = \max_h \left(\sum_{j=1}^N p_{ij}^h(r_{ij}^h + v_j(n)) \right).$$

A policy-iteration method involving a value-determination operation with a policy-improvement routine was given by R. A. Howard [4].

F. Dynamic Programming and the Calculus of Variations

Problems in the calculus of variations can be viewed as multistage decision problems of a continuous type. It can be shown that the dynamic-programming approach yields formal derivations of classical necessary conditions for the calculus of variations. Let us consider the problem of minimizing the functional

$$J(y) = \int_a^b f(x, y(x), y'(x)) dx,$$

where the function y is subject to $y(a) = c$.

We embed this problem within the family of problems generated by allowing a and c to be parameters with the ranges of variation $-\infty < a < b$, $-\infty < c < \infty$. Now we define the optimal value function $S(a, c) = \min_y J(y)$. Then the principle of optimality yields the functional equation

$$S(a, c) = \min \left(\int_a^{a+\Delta} f(x, y, y') dx + S(a+\Delta, c(y)) \right),$$

where the minimization is taken over all functions defined over $[a, a+\Delta]$ with $y(a) = c$ and $c(y) = y(a+\Delta)$. Then, writing $v = y'(a)$, we get

$$-\frac{\partial S}{\partial a} = \min_v \left(f(a, c, v) + v \frac{\partial S}{\partial c} \right).$$

This yields the Euler equation, the Legendre condition, the Weierstrass condition, and the Erdman corner conditions. Furthermore, it can be shown that the functional equation characterization yields the Hamilton-Jacobi partial differential equation of classical mechanics.

The dynamic-programming approach can be applied to more general problems in the calculus of variations.

G. Dynamic Programming and the Maximum Principle

In general, the method of dynamic programming carries a more universal character than the maximum principle of optimal control theory. However, in contrast to the latter, this

method does not have a rigorous logical foundation. V. G. Boltyanskii [6] has presented a justification of the dynamic programming method.

Let $f_i(x, u)$ ($i=0, 1, \dots, n$) be defined for $x \in V \subset \mathbf{R}^n$ and $u \in U \subset \mathbf{R}^r$, where V is an open set, and continuously differentiable on $V \times U$. Suppose that two points x^0 and x^1 are given in V . Among all the piecewise continuous controls $u(t) = (u_1(t), \dots, u_r(t)) \in U$ which transfer the phase point moving in accordance with

$$\frac{dx_i}{dt} = f_i(x, u(t)), \quad i=1, \dots, n,$$

from $x^0 = x(t_0)$ to $x^1 = x(t_1)$, find the control $u(t)$ for which the functional

$$J = \int_{t_0}^{t_1} f_0(x(t), u(t)) dt$$

takes the smallest value.

A continuous function $\omega(x) = \omega(x_1, \dots, x_n)$ is called a **Bellman function** relative to a point $a \in V$ if it possesses the following properties: (1) $\omega(a) = 0$; (2) there exists a set M (the singular set of $\omega(x)$), which is closed in V and does not contain interior points, such that the function $\omega(x)$ is continuously differentiable on the set $V - M$ and satisfies the condition

$$\sup_{u \in U} \left(\sum_{i=1}^n \frac{\partial \omega}{\partial x_i} f_i(x, u) - f_0(x, u) \right) = 0, \quad x \in V - M.$$

The following theorem gives a sufficient optimality condition.

Theorem: Assume that for $dx/dt = f(x, u(t))$ given in a region $V \subset \mathbf{R}^n$ there exists a Bellman function $\omega(x)$ relative to the point $a \in V$ with a piecewise smooth singular set. Assume, furthermore, that for any point $x^0 \in V$ there exists a control $u(t)$ which transfers the phase point from $x^0 = x(t_0)$ to $a = x(t_1)$ and satisfies the relation

$$\int_{t_0}^{t_1} f_0(x(t), u(t)) dt = -\omega(x^0).$$

Then any such control $u(t)$ is optimal in V .

Recently, Vinter and Levis [7] obtained the following general result in this connection. The sufficient condition is given in terms of a solution to the Bellman partial differential equation. It is shown that if this equation is modified so that it is actually an inequality, and if this inequality is required to be satisfied in a limiting sense only, then the condition is also necessary for optimality.

H. Characteristic Features of Dynamic Programming

The characteristic features of the dynamic-programming approach can be summarized in

the following five points: (1) the advantage of lower dimensionality in comparison with the enumeration approach; (2) the possibility of finding maxima and/or minima of functions defined over restricted domains for which differential calculus may not work well; (3) the availability of numerical solutions in recursive form; (4) the possibility of formulating certain problems to which classical methods do not apply; and (5) the applicability of the method to most types of problems in mathematical programming, such as inventory and production control, optimal searching, and some optimal and adaptive control processes.

References

- [1] R. E. Bellman, Dynamic programming, Princeton Univ. Press, 1957.
- [2] R. E. Bellman, Applied dynamic programming, Princeton Univ. Press, 1962.
- [3] R. E. Bellman, Adaptive control processes; a guided tour, Princeton Univ. Press, 1961.
- [4] R. A. Howard, Dynamic programming and Markov processes, Technology Press and Wiley, 1960.
- [5] S. E. Dreyfus, Dynamic programming and calculus of variations, Academic Press, 1965.
- [6] V. G. Boltyanskii, Sufficient conditions for optimality and the justification of the dynamic programming method, SIAM J. Control, 4 (1966), 326–361.
- [7] R. B. Vinter and R. M. Lewis, A necessary and sufficient condition for optimality of dynamic programming type, making no a priori assumption on the control, SIAM J. Control and Optimization, 16 (1978), 571–583.

E

128 (XVIII.15) Econometrics

A. General Remarks

The term **econometrics** can be interpreted in various ways. In its widest sense, it means the application of mathematical methods to economic problems and includes mathematical economics, †mathematical programming, etc. However, here we use it to mean statistical methods applied to economic analysis.

The object of econometrics is to provide methods to analyze relationships between economic variables. We classify these methods into four categories according to the types of relationships involved: (1) Analysis of causal relations: If a set of variables X_1, \dots, X_n affects an economic variable Y , we can estimate the direction and extent of those effects on Y . (2) Analysis of equilibrium: When a set of economic variables Y_1, \dots, Y_m is determined through a market equilibrium mechanism, we can analyze the structure of relationships that determines the equilibrium. (3) Analysis of correlation: When a set of economic variables is affected simultaneously by some (unknown) common factors, we can analyze the correlation structure of the variables. (4) Analysis of time interdependence: A process of development in time of a set of economic variables can be analyzed.

There are two types of economic data: (a) **macroeconomic data**, representing quantities and variables related to a national economy as a whole, usually based on national census; and (b) **microeconomic data**, representing information about the economic behavior of individual persons, households, and firms. Macroeconomic data are usually given as a set of time series (\rightarrow 421 Time Series Analysis A), while microeconomic data are obtained mainly through statistical surveys and are given as **cross-sectional data**. These two types of data, related to macroeconomic theory and microeconomic theory, respectively, require different approaches; and sometimes information obtained from both types of data has to be combined; obtaining macroeconomic information from microeconomic data is called aggregation.

B. Regression Analysis

The most common technique for the first category of problems is †regression analysis, which is applied to both microeconomic and macroeconomic analysis. However, there are

problems peculiar to economic analysis, where a factor can seldom be controlled; and usually there are too many highly related independent variables. In such cases, if all possible independent variables are taken into the model, the accuracy of the †estimators of the coefficients becomes extremely poor. Such a phenomenon, called **multicollinearity**, brings up the problem of selection of independent variables (\rightarrow 403 Statistical Models), to which no satisfactory solution has been given. Also, assumptions about the error terms may be dubious, the error terms may be correlated, or the variances may be different. If the †variance-covariance matrix of the errors is given, the †generalized least squares method can be applied, but usually such a matrix is not available.

C. Systems of Simultaneous Equations

The second category of problems is peculiar to economic analysis and applies mainly to macroeconomic data. Suppose that $\mathbf{Y} = (Y_1, \dots, Y_G)'$ is a vector consisting of G economic variables, among which there exist G relationships that determine the equilibrium levels of the variables. We also suppose that there exist K variables $\mathbf{Z} = (Z_1, \dots, Z_K)'$ that are independent of the economic relations but affect the equilibrium. The variables \mathbf{Y} are called **endogenous variables**, and the \mathbf{Z} are called **exogenous variables**. If we assume linear relationships among them, we have an expression such as

$$\mathbf{Y} = B\mathbf{Y} + \Gamma\mathbf{Z} + \mathbf{u}, \quad (1)$$

where B and Γ are matrices with constant coefficients and \mathbf{u} is a vector of **disturbances** or errors. (1) is called the **linear structural equation system** and is a system of simultaneous equations. By solving the equations formally, we get the so-called **reduced form**

$$\mathbf{Y} = \Pi\mathbf{Z} + \mathbf{v}, \quad (2)$$

where $\Pi = (I - B)^{-1}\Gamma$, $\mathbf{v} = (I - B)^{-1}\mathbf{u}$. The relation of \mathbf{Y} to \mathbf{Z} is determined through the reduced form (2), and if we have enough data on \mathbf{Y} and \mathbf{Z} we can estimate Π . The problem of **identification** is to decide whether we can determine the unknown parameters in B and Γ uniquely from the parameters in the reduced form. A necessary condition for the parameters in one of the equations in (1) to be identifiable is that the number of unknown parameters (or, since known constants in the system are usually set equal to zero, the number of variables appearing in the equation) not be greater than $K + 1$. If it is exactly equal to $K + 1$, the equation is said to be **just identified**,

and if it is less than $K + 1$, the equation is said to be **overidentified**.

If all the equations in the system are just identified, for arbitrary Π there exist unique B and Γ that satisfy $\Pi = (I - B)^{-1} \Gamma$. Therefore, if we denote the least squares estimator of Π by $\hat{\Pi}$, we can estimate B and Γ from the equation $(I - \hat{B})\hat{\Pi} = \hat{\Gamma}$. This procedure is called the **indirect least squares method** and is equivalent to the maximum likelihood method if we assume normality for u .

When some of the equations are overidentified, the estimation problem becomes complicated. Three kinds of procedures have been proposed: (1) full system methods, (2) single equation methods, and (3) subsystem methods. In full system methods all the parameters are considered simultaneously, and if normality is assumed, the maximum likelihood estimator can be obtained by minimizing $[(Y - \Pi Z)'(Y - \Pi Z)]$. Since it is usually difficult to compute the maximum likelihood estimator, a simpler, but asymptotically equivalent, **three-stage least squares method** has been proposed. The single equation methods and the subsystem methods take into consideration only the information about the parameters in one equation or in a subset of the equations, and estimate the parameters in each equation separately. There is a single equation method, called the **limited information maximum likelihood method**, based on the maximum likelihood approach, and also a **two-stage least squares method**, which estimates Π first by least squares, computes $\hat{Y} = \hat{\Pi}Z$, and then applies the least squares method to the model $Y = B\hat{Y} + \Gamma Z + \tilde{u}$.

These two and also some others are asymptotically equivalent. Among asymptotically equivalent classes of estimators corresponding to different information structures it has been established that the maximum likelihood estimators have asymptotically higher-order efficiency [5] (\rightarrow 399 Statistical Estimation) than other estimators, and Monte Carlo and numerical studies show that they are in most cases better than others if properly adjusted for the biases.

In many simultaneous equation models which have been applied to actual macroeconomic data, the values of endogenous variables obtained in the past appear on the right-hand sides of equations (1). Such variables are called **lagged variables**, and they can be treated, at least in the asymptotic theory of inference, as though they were exogenous. Hence exogenous variables and lagged endogenous variables are jointly called **predetermined variables**. When many lagged variables

appear over many time periods and when some structure among the coefficients of those lagged variables can be assumed, such a model is called a distributed lag model.

Sometimes it is necessary to include some nonlinear equations in the simultaneous equation model. Such nonlinear simultaneous equation models are difficult to deal with, partly because the solution of the equation may not be unique, and in practical applications ad hoc procedures are applied to obtain estimates of the parameters.

D. Multivariate and Time Series Analysis

Problems in the third category can be approached by multivariate analysis techniques. Sometimes principal component analysis and canonical correlation analysis have been applied to analyze the variations of a large amount of data. However, the practical meaning of the results obtained is often dubious.

The fourth category is the problem of time series analysis. Sophisticated theories of stochastic processes have little relevance for economic time series, because usually the time series do not satisfy such conditions as being stationary or having the Markov property, etc. Recently, however, autoregressive moving average (ARMA) and multivariate ARMA models [4] have been applied to macroeconomic data, especially for the purpose of prediction and for determining the direction of causal relations (\rightarrow 421 Time Series Analysis). Traditionally, fluctuations of economic time series have been thought to consist of trend, cyclic variation, seasonal variation, and error. Various ad hoc techniques have been used to separate or eliminate such components, but the theoretical treatment of such problems is far from satisfactory.

References

- [1] W. C. Hood and T. C. Koopmans (eds.), *Studies in econometric method*, Cowles Commission Monograph 14, Wiley, 1953.
- [2] H. Theil, *Economic forecasts and policy*, North-Holland, 1958.
- [3] J. Johnston, *Econometric methods*, McGraw-Hill, 1963.
- [4] G. E. P. Box and G. M. Jenkins, *Time series analysis: Forecasting and control*, Holden-Day, revised edition, 1976.
- [5] M. Akahira and K. Takeuchi, *Asymptotic efficiency of statistical estimators: Concepts and higher order efficiency*, Lecture notes in statistics 7, Springer, 1981.

129 (XXI.21) Einstein, Albert

Albert Einstein (March 14, 1879–April 18, 1955) was born of Jewish parents in the city of Ulm in southern Germany. He became a Swiss citizen soon after graduating from the Eidgenössische Technische Hochschule of Zürich in 1900. Afterward, he obtained a position as examiner of patents at Bern, and while at that post, he published his theories on light quanta, †special relativity, and †Brownian motion. After briefly holding professorships at the University of Zürich and the University of Prague, he became a professor at the University of Berlin in 1913. His general theory of relativity was announced in 1916, and in 1921 he won the Nobel Prize in physics for his contributions to theoretical physics. To escape Nazi persecution, he fled to the United States in 1933, and until his retirement in 1945 he was a professor at the Institute for Advanced Study at Princeton. He advised President Roosevelt of the feasibility of constructing the atomic bomb, but after World War II, along with others who had been connected with the bomb, he was active in promoting the nuclear disarmament movement and the establishment of a world government.

The theory of relativity raises fundamental epistemological problems concerning time, space, and matter. The results of the general theory were verified in 1919 by observations of the solar eclipse.

Through his latter years, Einstein continued to work on †unified field theory and on the generalization of relativity theory.

References

- [1] P. Frank, Einstein, his life and time, Knopf, 1947.
- [2] A. Einstein, The meaning of relativity, Princeton Univ. Press, fifth edition, 1956.
- [3] C. Seelig, Albert Einstein, eine dokumentarische Biographie, Europa Verlag, 1954.

130 (XX.16) Electromagnetism

A. Maxwell's Equations

Mathematical formulation of electromagnetics leads to †initial value and †boundary value problems for Maxwell's equations according to the geometric nature of the medium. **Max-**

well's equations for a vacuum are written in the form

$$\begin{aligned}\epsilon_0 \partial \mathbf{E} / \partial t &= \text{rot } \mathbf{H} - \mathbf{J}_e, & \epsilon_0 \text{div } \mathbf{E} &= \rho_e, \\ \mu_0 \partial \mathbf{H} / \partial t &= -\text{rot } \mathbf{E} - \mathbf{J}_m, & \mu_0 \text{div } \mathbf{H} &= \rho_m,\end{aligned}\quad (1)$$

where \mathbf{E} and \mathbf{H} are the **electric** and **magnetic field** vectors, ρ_e and ρ_m the electric and magnetic charge densities, \mathbf{J}_e and \mathbf{J}_m the electric and magnetic current densities, ϵ_0 and μ_0 constants, and the quantity $1/\sqrt{\epsilon_0 \mu_0} = c$ the speed of light in vacuum (2.99797×10^8 m/s). Charge and current densities must satisfy the **equations of continuity**

$$\begin{aligned}\partial \rho_e / \partial t + \text{div } \mathbf{J}_e &= 0, \\ \partial \rho_m / \partial t + \text{div } \mathbf{J}_m &= 0.\end{aligned}\quad (2)$$

Following upon the observation that, apparently, $\rho_m = 0$ and $\mathbf{J}_m = 0$ in nature, we henceforth set them equal to zero. This causes an asymmetry between the electric and magnetic quantities. On the other hand, the proposition " $\rho_m = 0$ and $\mathbf{J}_m = 0$ " cannot be deduced from the classical theory itself.

In the presence of matter, additional charge and current appear due to the **electric** and **magnetic polarization** \mathbf{P} and \mathbf{M} of the material. Therefore, in this case, it is necessary to make the following substitutions in (1):

$$\begin{aligned}\rho_e &\rightarrow \rho = \rho_e - \text{div } \mathbf{P}, \\ \mathbf{J}_e &\rightarrow \mathbf{J} = \mathbf{J}_e + \partial \mathbf{P} / \partial t + \text{rot } \mathbf{M}, \\ \mathbf{H} &\rightarrow \mathbf{H} + \mathbf{M}.\end{aligned}\quad (3)$$

Moreover, if we define the **electric flux density** (or **electric displacement**) \mathbf{D} and the **magnetic flux density** (or **magnetic induction**) \mathbf{B} by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}), \quad (4)$$

then Maxwell's equations (1) are transformed into

$$\begin{aligned}\partial \mathbf{D} / \partial t &= \text{rot } \mathbf{H} - \mathbf{J}_e, & \text{div } \mathbf{D} &= \rho_e, \\ \partial \mathbf{B} / \partial t &= -\text{rot } \mathbf{E}, & \text{div } \mathbf{B} &= 0.\end{aligned}\quad (5)$$

In the electromagnetic field in a vacuum there is energy with a density

$$u = (\epsilon_0/2) \mathbf{E}^2 + (\mu_0/2) \mathbf{B}^2, \quad (6)$$

and energy flux with a density expressed by the **Poynting vector**

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}. \quad (7)$$

Between these quantities the following relation holds:

$$\partial u / \partial t + \text{div } \mathbf{S} = 0. \quad (8)$$

An electric charge q moving with velocity \mathbf{v} in an electromagnetic field is subject to the **Lorentz force**

$$\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}. \quad (9)$$

This force can be interpreted as being caused by the **Maxwell stress tensor**

$$T_{ik} = (\epsilon_0/2)(-E_i E_k + 2\delta_{ik} E^2) + (\mu_0/2)(-H_i H_k + 2\delta_{ik} H^2). \quad (10)$$

By introducing the **scalar potential** V and the **vector potential** \mathbf{A} , we can express the field vectors as

$$\mathbf{B} = \text{rot } \mathbf{A}, \quad \mathbf{E} = -\text{grad } V - \partial \mathbf{A} / \partial t. \quad (11)$$

Furthermore, if we impose an auxiliary condition (**Lorentz condition**)

$$\text{div } \mathbf{A} + \epsilon_0 \mu_0 \partial V / \partial t = 0, \quad (12)$$

then we obtain from (1) the wave equations

$$\square V = -\rho_e / \epsilon_0, \quad \square \mathbf{A} = -\mu_0 \mathbf{J}_e, \quad (13)$$

where $\square \equiv \Delta - \epsilon_0 \mu_0 \partial^2 / \partial t^2$ is called the **d'Alembertian** and is sometimes written as \diamond^2 . From (13) we conclude that the electromagnetic field can propagate in a vacuum as a wave with speed $c = 1/\sqrt{\epsilon_0 \mu_0}$. In \dagger quantum theory the potentials V and \mathbf{A} are regarded as being more fundamental than \mathbf{E} and \mathbf{H} themselves. However, they are not uniquely determined, in the sense that the **gauge transformation**

$$V \rightarrow V + \partial \psi / \partial t, \quad \mathbf{A} \rightarrow \mathbf{A} - \text{grad } \psi \quad (14)$$

with an arbitrary function ψ of the space and time variables does not affect the fields.

Maxwell's equations are invariant under the \dagger Lorentz transformation. Therefore they can be written in 4-dimensional tensor form (\rightarrow 359 Relativity C).

Maxwell's equations can be regarded as the \dagger wave equations for \dagger bosons with spin 1 (\dagger photons). The equations of quantum electrodynamics are obtained if we regard the field quantities as quantum-mechanical variables (q -numbers) and then perform \dagger second quantization.

B. Concrete Problems

In solving Maxwell's equations concretely, we usually make additional assumptions for polarizations, electric current, and field vectors

$$\mathbf{P} = \chi_e \mathbf{E}, \quad \mathbf{M} = \chi_m \mathbf{H}, \quad \mathbf{J}_c = \sigma \mathbf{E}, \quad (15)$$

where χ_e , χ_m , and σ are called the **electric susceptibility**, **magnetic susceptibility**, and **conductivity**, respectively. Then we have

$$\mathbf{D} = \epsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}, \quad (16)$$

where $\epsilon = \epsilon_0 + \chi_e$ is the **dielectric constant** and $\mu = \mu_0(1 + \chi_m)$ is the **magnetic permeability**. Therefore the equations (5) become identical to (1) if ϵ_0 and μ_0 are replaced by ϵ and μ , respectively. (ρ_m and \mathbf{J}_m are set equal to zero.)

Some cases of practical importance are given below.

(1) **Electrostatics**. If the fields are time-independent and there is no electric current, then \mathbf{E} and \mathbf{H} are mutually independent. The static electric field is calculated from the solution of the boundary value problem of the \dagger Poisson equation $\Delta V = -\rho_e/\epsilon$. Specifically, V takes a constant value in each conductor.

(2) **Magnetostatics**. For zero electric current, the problem of magnetostatics is solved in the same way as in electrostatics. For the case of nonvanishing stationary electric current, the problem is reduced to that of solving

$$\Delta \mathbf{A} = -\mu \mathbf{J}_e, \quad \text{div } \mathbf{A} = 0. \quad (17)$$

(3) **Electric current in a conductor**. A stationary electric current in a conductor is governed by the equation of continuity $\text{div } \mathbf{J} = 0$, **Ohm's law** $\mathbf{J} = \sigma \mathbf{E}$, and a special case of Maxwell's equation $\text{rot } \mathbf{E} = 0$. The electric current produces heat (**Joule heat**) proportional to $\mathbf{J} \cdot \mathbf{E}$ per unit volume per unit time (**Joule's law**). For certain substances, the **specific resistance** σ^{-1} suddenly becomes negligibly small below a critical temperature. This is called **superconductivity**.

(4) **Quasistationary electric circuit**. The problem appearing most often in electrical engineering is that of a quasistationary circuit. Its characteristic feature is that the electric currents exist only in the circuit elements (inductors, capacitors, and resistors) and in the lines connecting them. The current (both \mathbf{J}_e and $\partial \mathbf{D} / \partial t$) can be neglected in all other parts of the system. (This could be compared to the situation in dynamics where we consider systems of material points or of rigid bodies having a finite number of degrees of freedom, although every material body is essentially a continuum.) The system network is constructed as a \dagger linear graph with the circuit elements as its branches. Topological \dagger network theory deals with the relation between the structure of the linear graph and the electrical characteristics of the network, whereas function-theoretic network theory deals with the relation between current and voltage at each part of the network. In the latter theory, current and voltage are considered as functions of the frequency of the sinusoidal alternating voltage applied to some point of the network. Together these constitute a unique theoretical system in engineering mathematics for designing system networks (\rightarrow 282 Networks).

(5) **Theory of electromagnetic waves**. The theory of electromagnetic waves deals with the case where the changes of all field quantities are proportional to $e^{i\omega t}$, and in addition the frequency ω is so large that the term $\partial \mathbf{D} / \partial t$ in

(5) is of the same order as or larger than \mathbf{J}_e . In such a situation the electromagnetic field behaves like a wave.

Problems of various types arise depending on the geometry of the conducting and dielectric substances, on the type of the energy source, etc. Important problems are: (i) radiation of a wave from a point source into free space; (ii) scattering of a plane wave by small bodies or cylinders; (iii) diffraction of a wave through holes in a conducting plate; (iv) reflection and refraction of a wave at the boundary between different media; (v) wave propagation along a conducting tube (wave guide); and (vi) resonance of the electromagnetic field in a cavity surrounded by a conducting substance. Theoretical treatment similar to that for ordinary networks is possible for microwave circuits consisting of wave guides, cavities, etc.

(6) **Wave guides.** For electromagnetic waves with a harmonic time dependence $e^{-i\omega t}$ propagating inside a hollow tube of uniform cross section extending in the z -direction with perfectly conducting walls (called a **wave guide**), the Maxwell equations become

$$\begin{aligned}\operatorname{rot} \mathbf{E} &= i\omega \mathbf{B}, & \operatorname{div} \mathbf{B} &= 0, \\ \operatorname{rot} \mathbf{B} &= -i\mu\epsilon\omega \mathbf{E}, & \operatorname{div} \mathbf{E} &= 0,\end{aligned}\quad (18)$$

with the boundary condition

$$\mathbf{n} \times \mathbf{E} = 0, \quad \mathbf{n} \cdot \mathbf{B} = 0,$$

where \mathbf{n} is a unit normal at the boundary surface. A further reduction is gained by Fourier analysis in the z -variable. For harmonic z -dependence e^{ikz} , the transverse components $\mathbf{E}_t = (\mathbf{e}_z \times \mathbf{E}) \times \mathbf{e}_z$ and $\mathbf{B}_t = (\mathbf{e}_z \times \mathbf{B}) \times \mathbf{e}_z$ are determined from the z -components $E_z = \mathbf{e}_z \cdot \mathbf{E}$ and $H_z = \mathbf{e}_z \cdot \mathbf{H}$ by the following part of the Maxwell equations:

$$\begin{aligned}ikE_t + i\omega\mathbf{e}_z \times \mathbf{B}_t &= \operatorname{grad}_t E_z, \\ ik\mathbf{B}_t - i\mu\epsilon\omega\mathbf{e}_z \times \mathbf{E}_t &= \operatorname{grad}_t B_z\end{aligned}\quad (19)$$

(where grad_t is the transverse components of the gradient), up to the solutions for $E_z = B_z = 0$, called **transverse electromagnetic (TEM) waves**, for which $k = \omega\sqrt{\mu\epsilon}$, $\mathbf{B}_t = \pm\sqrt{\mu\epsilon}\mathbf{e}_z \times \mathbf{E}_t$, and \mathbf{E}_t is a solution of the electrostatic problem in two dimensions $\operatorname{rot} \mathbf{E}_t = 0$, $\operatorname{div} \mathbf{E}_t = 0$. To have $\mathbf{E}_t \neq 0$ for the TEM solution, it is necessary to have two or more surfaces, such as a coaxial cable (region between two concentric cylinders) or a parallel-wire transmission line. Nonzero longitudinal components E_z and B_z are determined from the 2-dimensional equations

$$\left[\left(\frac{\partial}{\partial x} \right)^2 + \left(\frac{\partial}{\partial y} \right)^2 + (\mu\epsilon\omega^2 - k^2) \right] \begin{Bmatrix} E_z \\ B_z \end{Bmatrix} = 0 \quad (20)$$

with boundary conditions $E_z = 0$ and $\partial B_z / \partial n = 0$ on the wall, where $(\partial / \partial n)$ denotes the partial derivative in the normal direction. The solutions with $B_z = 0$ are called **transverse magnetic (TM) waves** (or **electric (E) waves**) and those with $E_z = 0$ **transverse electric (TE) waves** (or **magnetic (M) waves**); for each of these cases the equations determine an eigen-wave number k for a given angular frequency ω (typically in a waveguide situation) or eigenfrequency $\nu = \omega / (2\pi)$ for a given wavelength $\lambda = 2\pi / k$ (typically in a resonant cavity situation).

References

- [1] J. D. Jackson, *Classical electrodynamics*, Wiley, second edition, 1975.
- [2] W. K. H. Panofsky and M. Phillips, *Classical electricity and magnetism*, Addison-Wesley, second edition, 1961.
- [3] A. Sommerfeld, *Electrodynamics*, trans. by E. G. Ramberg, Academic Press, 1952. (Original in German, 1948.)
- [4] P. Moon and D. E. Spencer, *Foundations of electrodynamics*, Van Nostrand, 1960.
- [5] J. A. Stratton, *Electromagnetic theory*, McGraw-Hill, 1941.
- [6] B. I. Bleaney, *Electricity and magnetism*, Clarendon Press, 1957.
- [7] R. Fano, L. J. Chu, and R. B. Adler, *Electromagnetic fields, energy and forces*, Wiley, 1960.
- [8] E. G. Hallen, *Electromagnetic theory*, Chapman & Hall, 1962.

131 (X.8) Elementary Functions

A. Definition

A function of a finite number of real or complex variables that is algebraic, exponential, logarithmic, trigonometric, or inverse trigonometric, or the composite of a finite number of these, is called an **elementary function**. Elementary functions comprise the most common type of function in elementary calculus.

J. Liouville [1] defined the elementary functions as follows: An algebraic function of a finite number of complex variables is called an elementary function of class 0. Then e^z and $\log z$ are called elementary functions of class 1. Inductively, we define elementary functions of class n under the assumption that elementary functions of class at most $n - 1$ have already

been defined. Let $g(t)$ and $g_j(w_1, \dots, w_n)$ ($1 \leq j \leq m$) be elementary functions of class at most 1 and $f(z_1, \dots, z_m)$ be an elementary function of class at most $n-1$. Then the composite functions $g(f(z_1, \dots, z_m))$ and $f(g_1(w_1, \dots, w_n), \dots, g_m(w_1, \dots, w_n))$ (and only such functions) are called the elementary functions of class at most n . An elementary function of class at most n and not of class at most $n-1$ is called an **elementary function of class n** . A function that is an elementary function of class n for some integer n is called an **elementary function**. In this article, we explain the properties of the most common elementary functions.

B. Exponential and Logarithmic Functions of a Real Variable

Let $a > 0, a \neq 1$. A function $f(x)$ of a real variable satisfying the functional relation

$$f(x+y) = f(x)f(y), \quad f(1) = a, \tag{1}$$

satisfies $f(n) = a^n$ for positive integers n and $f(-n) = 1/a^n$ for negative integers $-n$. In general, $f(n/m) = \sqrt[m]{a^n}$ for every rational number $r = n/m$. If we assume that $f(x)$ is continuous, then there is a unique strictly monotone function $f(x)$ defined in $(-\infty, \infty)$ whose range is $(0, \infty)$. The function $f(x)$ is called the **exponential function with the base a** and is denoted by a^x , read “ a to the power x ” and also called a **power of a with exponent x** . Its inverse function is called the **logarithmic function to the base a** , and is denoted by $\log_a x$. The specific value $\log_a x$ is called the **logarithm of x to the base a** . If $g(x) = \log_a x$, we have

$$g(xy) = g(x) + g(y), \quad g(a) = 1. \tag{2}$$

Hence we have $xy = f(g(x) + g(y))$. Therefore we can reduce multiplication to addition by using a numerical table for the logarithmic function.

C. Logarithmic Computation

The logarithm to the base 10 is called the **common logarithm**. If two numbers x, y expressed in the decimal system differ only in the position of the decimal point (i.e., $y = x \cdot 10^n$ for an integer n), they share the same fractional parts in their common logarithms. The integral part of the common logarithm is called the **characteristic**, and the fractional part is called the **mantissa**. (We note that the word “mantissa” is also frequently used for the fractional part a in the floating point representation $x = a \cdot 10^n, 10^{-1} \leq a < 1$, or $1 \leq a < 10$.) The

common logarithms of integers have been computed and published in tables.

D. Derivatives of Exponential and Logarithmic Functions

The function $f(x) = a^x$ is differentiable, and $f'(x) = k_a f(x)$, where k_a is a constant determined by the base a . If we take the base a to be

$$e = \lim_{v \rightarrow \infty} \left(1 + \frac{1}{v}\right)^v = \sum_{v=0}^{\infty} \frac{1}{v!} = 2.71828 \dots,$$

then we have $k_e = 1$. The number $\sum(1/v!)$ is usually called **Napier’s number** and is denoted by e after L. Euler (his letter to C. Goldbach of 1731; → Appendix B, Table 6). In 1873, C. Hermite proved that e is a transcendental number. We sometimes denote e^x by **exp x** ; the term **exponential function** usually means the function $\exp x$. The function e^x is invariant under differentiation, and conversely, a function invariant under differentiation necessarily has the form Ce^x . The logarithm to the base e is called the **Napierian logarithm (or natural logarithm)**, and we usually denote it by **log x** without explicitly naming the base e (sometimes it is denoted by $\ln x$). The derivative of $\log x$ is $1/x$, hence we have the integral representation

$$\log x = \int_1^x \frac{dx}{x}. \tag{3}$$

The constant factor k_a in the derivative of a^x is equal to $\log a$. The graphs of $y = e^x$ and $y = \log x$ are shown in Fig. 1. The functions e^x and $\log(1+x)$ are expanded in the following Taylor series at $x=0$:

$$e^x = \sum_{v=0}^{\infty} \frac{x^v}{v!}, \tag{4}$$

$$\log(1+x) = \sum_{v=1}^{\infty} (-1)^{v+1} \frac{x^v}{v}. \tag{5}$$

The power series in the right-hand sides of (4) and (5) are called the **exponential series** and the **logarithmic series**, respectively. The radii of convergence of (4) and (5) are ∞ and 1, respectively.

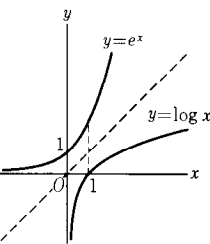


Fig. 1

E. Trigonometric and Inverse Trigonometric Functions of a Real Variable

The trigonometric functions of a real variable x are the functions $\sin x$, $\cos x$ (\rightarrow 432 Trigonometry) and the functions $\tan x = \sin x / \cos x$, $\cot x = \cos x / \sin x$, $\sec x = 1 / \cos x$, and $\operatorname{cosec} x = 1 / \sin x$ derived from $\sin x$ and $\cos x$. The derivatives of $\sin x$ and $\cos x$ are $\cos x$ and $-\sin x$, respectively. They have the following Taylor expansions at $x=0$:

$$\sin x = \sum_{v=0}^{\infty} \frac{(-1)^v}{(2v+1)!} x^{2v+1}, \quad (6)$$

$$\cos x = \sum_{v=0}^{\infty} \frac{(-1)^v}{(2v)!} x^{2v}. \quad (7)$$

The radii of convergence of (6) and (7) are both ∞ .

The inverse functions of $\sin x$, $\cos x$, and $\tan x$ are the **inverse trigonometric functions** and are denoted by $\arcsin x$, $\arccos x$, and $\arctan x$, respectively. (Instead of this notation, $\sin^{-1} x$, $\cos^{-1} x$, and $\tan^{-1} x$ are also used). These functions are infinitely multiple-valued, as shown in Fig. 2. But if we restrict their ranges within the part shown by solid lines in Fig. 2, they are considered single-valued functions. To be more precise, we restrict the range as follows: $-\pi/2 \leq \arcsin x \leq \pi/2$, $0 \leq \arccos x \leq \pi$, $-\pi/2 < \arctan x < \pi/2$.

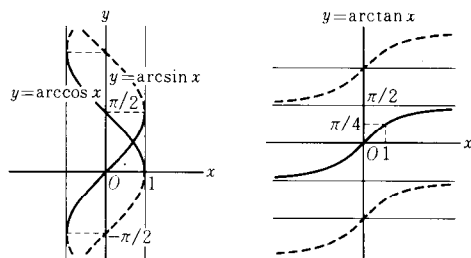


Fig. 2

The functions having these ranges are called the **principal values** and are sometimes denoted by $\operatorname{Arcsin} x$, $\operatorname{Arccos} x$, and $\operatorname{Arctan} x$, respectively. The derivatives of these functions are $(1-x^2)^{-1/2}$, $-(1-x^2)^{-1/2}$, $(1+x^2)^{-1}$, respectively (\rightarrow Appendix A, Table 9.I; for the Taylor or Laurent expansions of $\tan x$, $\cot x$, $\sec x$, $\operatorname{cosec} x$, $\arcsin x$, $\arccos x$, $\arctan x$, etc., \rightarrow Appendix A, Table 10.IV).

F. Hyperbolic Functions

Let P be a point on the branch of the hyperbola $x^2 - y^2 = 1$, $x > 0$, and let O be the origin and A the vertex $(1, 0)$ of the hyperbola. Denote by $\theta/2$ the area of the domain surrounded by the line segments OA , OP , and the arc \widehat{AP}

of the hyperbola. Then we define the coordinates of P to be $(\cosh \theta, \sinh \theta)$ as functions of θ . We have

$$\begin{aligned} \cosh x &= (e^x + e^{-x})/2, \\ \sinh x &= (e^x - e^{-x})/2, \end{aligned} \quad (8)$$

called the **hyperbolic cosine** and **hyperbolic sine**, respectively. As in the case of trigonometric functions, we define the **hyperbolic tangent** by $\tanh x = \sinh x / \cosh x$, the **hyperbolic cotangent** by $\coth x = \cosh x / \sinh x$, the **hyperbolic secant** by $\operatorname{sech} x = 1 / \cosh x$, and the **hyperbolic cosecant** by $\operatorname{cosech} x = 1 / \sinh x$. They are called the **hyperbolic functions**. The graphs of $\sinh x$ and $\cosh x$ are shown in Fig. 3. The trigonometric functions are sometimes called **circular functions**.

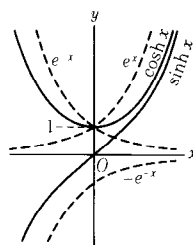


Fig. 3

We now introduce the **Gudermannian** (or **Gudermann function**):

$$\theta = \operatorname{gd} u = 2 \arctan e^u - \pi/2,$$

$$u = \operatorname{gd}^{-1} \theta = \log |\tan \theta + \sec \theta| = \frac{1}{2} \log \frac{1 + \sin \theta}{1 - \sin \theta}.$$

Then the hyperbolic functions can be expressed in terms of the trigonometric functions. For example,

$$\sinh u = \tan \theta, \quad \cosh u = \sec \theta, \quad \tanh u = \sin \theta.$$

G. Elementary Functions of a Complex Variable (\rightarrow Appendix A, Table 10)

(1) **Exponential function.** The power series (4) converges for all finite values if we replace x by the complex variable z and gives an entire function of z with an essential singularity at the point at infinity. This is the exponential function e^z of a complex variable z . It satisfies the **addition formula** (1), $e^{z_1+z_2} = e^{z_1} e^{z_2}$, and it is also the analytic continuation of the exponential function of a real variable. For a purely imaginary number $z = iy$, we have the **Euler formula**

$$e^{iy} = \cos y + i \sin y. \quad (9)$$

The function $w = e^z$ gives a conformal mapping from the z -plane to the w -plane, as shown in Fig. 4, which maps the imaginary axis of

the z -plane onto the unit circle of the w -plane ($w = u + iv$). For $z = x + iy$ (x, y are real numbers), we have $e^z = e^x e^{iy}$; hence e^z is a simply periodic function with fundamental period $2\pi i$.

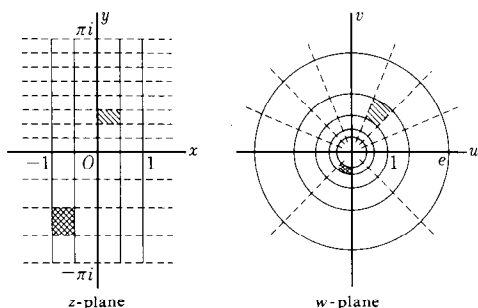


Fig. 4

(2) **Logarithmic function.** The logarithmic function $\log z$ of a complex variable z is the inverse function of e^z . It is an infinitely multiple-valued analytic function that has logarithmic singularities at $z = 0$ and $z = \infty$. All possible values are expressed by $\log z + 2n\pi i$ (n is an arbitrary integer), where we select a suitable value $\log z$. The **principal value** of $\log z$ is usually taken as $\log r + i\theta$, where $z = re^{i\theta}$ ($r = |z|$, θ is the argument of z) and $0 \leq \theta < 2\pi$. (Sometimes the range of the argument is taken as $-\pi < \theta \leq \pi$.) The principal value of $\log z$ is sometimes denoted by **Log** z . The power series (5) gives one of its functional elements. The integral representation (3) holds for a complex variable z . The multivalency of $\log z$ results from the selection of a contour of integration; the integral of $1/z$ around the origin is $2\pi i$, which is the increment of $\log z$.

(3) **Power.** The exponential function a^z for an arbitrary complex number a is defined to be $\exp(z \log a)$. Similarly, z^a is defined to be $\exp(a \log z)$. The function z^a is an algebraic function if and only if a is rational. In other cases, the function z^a is an elementary function of class 2.

(4) **Trigonometric functions, inverse trigonometric functions, hyperbolic functions.** The trigonometric, inverse trigonometric, and hyperbolic functions of a complex variable are defined by the analytic continuations of the corresponding functions of a real variable. For example, $\sin z$ and $\cos z$ are defined by the power series (6) and (7), respectively. They are entire functions whose zero points are $n\pi$ and $(n - \frac{1}{2})\pi$ (n is an integer), respectively. They are also represented by Weierstrass's infinite product (\rightarrow Appendix A, Table 10.VI).

The functions $\tan z$, $\cot z$, $\sec z$, and $\operatorname{cosec} z$ are meromorphic functions of z on the complex z -plane, and they are expressed by Mittag-Leffler partial fractions (\rightarrow Appendix A, Tables 10.IV, 10.V). As can be shown from

(8) and (9), we have

$$\begin{aligned} \cos z &= (e^{iz} + e^{-iz})/2, & \sin z &= (e^{iz} - e^{-iz})/2i, \\ \cosh z &= \cos iz, & \sinh z &= (\sin iz)/i. \end{aligned} \quad (10)$$

Each of these formulas (10) is called an **Euler formula**. For a complex variable, the trigonometric and hyperbolic functions are composites of exponential functions, the inverse trigonometric functions are composites of logarithmic functions, and all of them are elementary functions of class 1. The definition of elementary functions by Liouville described in Section A refers, of course, to the functions of a complex variable. We remark that the inverse function of an elementary function is not necessarily an elementary function. For example, the inverse function of $y = x - a \sin x$ is not an elementary function (\rightarrow 309 Orbit Determination B).

The derivative of an elementary function is also an elementary function. However, the primitive function of an elementary function is not necessarily an elementary function. The primitive function of a rational function or an algebraic function of genus 0 is again an elementary function. Similar properties hold for rational functions of trigonometric functions. Liouville [1] carried through a deep investigation of the situation where the integral of an elementary function is also an elementary function.

References

- [1] J. Liouville, Mémoire sur la classification des transcendentes et sur l'impossibilité d'exprimer les racines de certaines équations en fonction finie explicite des coefficients, J. Math. Pures Appl., 2 (1837), 56–104; 3 (1838), 523–546.
 - [2] E. G. H. Landau, Einführung in die Differentialrechnung und Integralrechnung, Noordhoff, 1934; English translation: Differential and integral calculus, Chelsea, 1965.
 - [3] N. Bourbaki, Eléments de mathématique, Fonctions d'une variable réelle, ch. 3, Fonctions élémentaires, Actualités Sci. Ind., 1074b, Hermann, second edition, 1958.
- Also \rightarrow references to 106 Differential Calculus.

132 (XX.31) Elementary Particles

A. Introduction

The word “atom” is derived from the Greek word for indivisible. It turns out that an atom

is divisible into its constituent nucleus and electrons. The nucleus, in turn, consists of protons and neutrons (together called nucleons). Photons (quanta of electromagnetic waves), electrons, protons, and neutrons (denoted by γ , e , p , and n , respectively), along with many other subsequently discovered sub-nuclear particles, are called **elementary particles**, while nuclei, atoms, and molecules are the **composite particles** composed of these elementary particles.

States of a particle form an irreducible (unitary) representation of the proper inhomogeneous Lorentz group with positive energy (or a finite direct sum of such representations). Thus a **mass** ($m \geq 0$) and a **spin** ($j = 0, \frac{1}{2}, 1, \dots$) are assigned to each particle (\rightarrow 258 Lorentz Group). For example, e , p , and n have spin $\frac{1}{2}$ and nonzero masses, while γ has spin 1 and zero mass.

Many elementary particles are unstable, decaying into other particles. The average **lifetime** is denoted by τ and its inverse is called the **half-width**. The time in which half of many samples of the same particle decays is called the **half-life**, given by $\tau \log 2$. For example, electrons, protons, and photons are supposed to be stable (or at least to have very long lifetimes), while a neutron decays into a proton, an electron, and a neutrino ν (β -decay) with a lifetime of about 15 minutes.

From the study of the relativistic equations (Dirac equations) for wave functions of an electron, P. A. M. Dirac predicted the existence of particles with the same mass as the electron, but of opposite electric charge (Dirac's hole theory, 1930). These were discovered in 1932 and called positrons. Every elementary particle is now believed to be associated with an **anti-particle** characterized by the opposite sign of the particle's additive quantum numbers, the two being connected by the \dagger PCT theorem. Hence the positron is the antiparticle of the electron. The antiproton, theoretically expected for a long time and experimentally found in 1955, is the antiparticle of the proton. Antiprotons, antineutrons, and positrons are constituents of antimatter. The particles whose additive quantum numbers are all invariant under change of sign, such as photons (γ), neutral pions (π^0), etc., seem to be antiparticles of themselves and are said to be self-conjugate.

Elementary particles have four distinct types of interaction: gravitational, weak, electromagnetic, and strong, in increasing order of interaction strength. Gravitational and electromagnetic interactions were recognized in earlier centuries because these interactions are of long range. A. Einstein put forward the idea of a light quantum or photon as a lump of electromagnetic energy behaving like a par-

ticle. Whether a corresponding quantum (graviton) exists for gravitational interactions is a question related to the existence of gravitational waves themselves, and is not yet settled.

The nuclear force is an example of a strong interaction and is studied to elucidate nuclear structure and to derive the cross sections of various collision processes involving nuclei. H. Yukawa predicted in 1935 the existence of a particle associated with the nuclear force, just as photons are associated with the electromagnetic interaction. Its mass was predicted, from the range of the nuclear force, to be about 200 times the electron mass, which is intermediate between the masses of electrons and nucleons, and hence the particle was named a mesotron or a meson. These were found in cosmic rays in 1947 and, in fact, it was found that the mesons relevant to the nuclear force (now called pions and denoted by π^+ , π^- , π^0 according to their electric charge) decay into other kinds of particles called muons (denoted by μ^+ , μ^-) with a charged-pion lifetime of 2.6×10^{-8} sec ($\pi^\pm \rightarrow \mu^\pm + \nu(\bar{\nu})$; the neutral pion π^0 decays, over a shorter lifetime, into photons). The muons then decay into electrons and neutrinos ($\mu^\pm \rightarrow e^\pm + \nu + \bar{\nu}$, $\bar{\nu}$ indicating antineutrinos) with muon lifetimes of about 2.2×10^{-6} sec. Weak interactions are relevant to these decays as well as to the β -decay of the nucleus and the neutron. Since 1962 electron neutrinos ν_e and muon neutrinos ν_μ have been distinguished in these decays, so that electron and muon numbers may be conserved.

Since 1949, many new particles (unstable under weak interactions) have been gradually found in cosmic rays; these are called strange particles. Some of the early ones are hyperons (Λ , Σ , Ξ), of spin $\frac{1}{2}$, and kaons (K , \bar{K}), of spin 0. For such strange particles, the strangeness quantum number, which is preserved in the strong interaction, has been introduced, and the **Nakano–Nishijima–Gell-Mann formula** concerning this number is known to hold. This says that $Q = I_3 + \frac{1}{2}(B + S)$ for each elementary particle, where Q is the electric charge in units of that of the positron, I_3 is the third component of the isospin, B is the baryon number, and S is the strangeness.

Quantum-mechanical wave functions for a system of identical particles seem to be either totally symmetric (Bose statistics) or totally antisymmetric (Fermi statistics) under permutations of particles. Accordingly, particles are either **bosons** (or Bose particles) or **fermions** (or Fermi particles). All bosons seems to have integer spins and all fermions half-odd-integer spins. This is the **connection of spin and statistics** and follows from certain axioms in quantum field theory (\rightarrow 150 Field Theory). There

has been some discussion on intermediate statistics (parabosons and parafermions).

B. Families of Elementary Particles

Elementary particles are classified into families of leptons, photons, and hadrons.

The family of leptons now consists of electrons (e), muons (μ), and tau-leptons (τ), their accompanying neutrinos (ν_e, ν_μ, ν_τ), and their antiparticles; τ was discovered in the late 1970s. Experimentally, ν_τ is not well established nor has the possibility $\nu_\tau = \nu_e$ yet been excluded. Leptons have spin $\frac{1}{2}$. They are characterized by having no strong interactions.

The family of photons consists of photons and the recently discovered intermediary weak vector bosons. Gluons (\rightarrow Section C (4)), if they exist, also belong to this family.

The family of hadrons has a large number of members which are either **mesons** or **baryons**, the former being bosons and the latter fermions. We now have, in addition to pions and kaons, many resonant mesons (unstable under strong interaction) such as ρ -mesons and ω -mesons. At present we have, in all, more than 20 species of mesons. This number does not count spin, charge, and antiparticle degrees of freedom. The baryon subfamily includes nucleons, hyperons, and excited states, now consisting of more than 30 species, again not counting spin, charge, and antiparticle degrees of freedom. We have nucleonic resonances with spin as high as $11/2$.

Hadrons are now well understood as composites of subhadronic constituents called quarks (and antiquarks), although free quarks have not been observed. (Hence the problem of quark confinement has been discussed extensively.) Mesons are systems made up of a quark and an antiquark. Baryons are systems made up of three quarks. Therefore, at our present level of knowledge, the elementary particles might be leptons, photons, and quarks (and the corresponding antiparticles).

C. Methods in the Theory of Elementary Particles

(1) **†Quantum Field Theory.** Application of the ideas of quantum mechanics to electromagnetic fields and their interaction with electrons resulted in the formulation of **quantum electrodynamics** (and more generally quantum field theory). Application of quantum-mechanical perturbation theory to quantum electrodynamics with the fine-structure constant $\alpha = e^2/(\hbar c)$ (about $1/137$) as an expansion parameter resulted in divergent expressions—

the so-called divergence difficulties. The **ultra-violet divergence** comes from integration over high momenta (of virtual particles), and the **infrared divergence** is due to the zero mass of photons. It was later found that the ultraviolet divergence can be combined with a small number of parameters of the theory (i.e., electron mass, (possibly nonzero) photon mass, and electromagnetic coupling constant e) into a revised set of constants (called the renormalized masses and coupling constant), which are then equated to the observed finite values of these constants—a procedure called renormalization. Physically, this renormalization is pictured to be effected by virtual photons (and electron-positron pairs) surrounding (bare) electrons and photons. The infrared divergence is supposed to be a reflection of the fact that electrons can be accompanied by infinitely many photons with negligibly small total energy (a situation made possible by the zero mass of photons); this cannot be experimentally analyzed (and is indistinguishable from a single electron).

The relativistically covariant formulation of the renormalized perturbation theory of quantum electrodynamics proposed by S. Tomonaga, J. S. Schwinger, and R. P. Feynman (independently and in different forms), and in particular the Feynman rules and Feynman diagrams that lead to the Feynman integrals (\rightarrow 146 Feynman Integrals), made possible detailed theoretical computations, and the computed values (such as the Lamb shift of hydrogen and the anomalous magnetic moment of an electron) fit marvelously well with observed values—an achievement considered a great success of quantum electrodynamics. F. J. Dyson more or less showed that the renormalization procedure really absorbs all the divergences in all orders of the perturbation expansion in terms of renormalized constants, though there were later refinements and elaborations of the proof. This work also leads to the division of quantum field theories into two classes: **renormalizable** theories, where (infinitely many) ultraviolet divergences can be absorbed into a finite number of constants by renormalized perturbation theory, and **unrenormalizable** theories. The question of whether perturbation series converge in some sense is an unsolved question of quantum electrodynamics.

In quantum field theory, the central role is played by quantum fields, which are operator-valued generalized functions of a space-time point. Particle interpretations of any state at infinite past and infinite future are obtained in the theory from the asymptotic behavior of the fields (\dagger asymptotic fields) at time $-\infty$ and time $+\infty$, and from their relation, expressed by the

\dagger S-matrix, describing how particles scatter by collision. Thus any model of quantum fields makes, in principle, a prediction about what particles appear and how they behave (asymptotically) in mutual collisions.

After the success of quantum electrodynamics, the perturbation theory of quantum fields was applied to systems of pions and nucleons; this proved to be unsuccessful, possibly due to the lack of an appropriate expansion parameter. This has led some people to study the mathematically rigorous consequences of quantum field theories; these mathematical consequences do not require any perturbation calculation, simply following from a small number of mathematically formulated axioms believed to be satisfied by a large class of quantum field theories. This approach, referred to as \dagger axiomatic quantum field theory, has yielded a few physically meaningful consequences of general nature: analyticity of some S -matrix elements, the \dagger PCT theorem, and the connection of spin and statistics (\rightarrow Section A).

While the axiomatic approach provides a general framework, the \dagger constructive field theory developed later provides examples of quantum field theories that fit into such a framework. Because of its concrete nature, it can make statements about detailed properties of the model, such as the (non-)existence of composite particles, the establishment of perturbation theory as an asymptotic expansion, and phase-transitions phenomena and the related broken symmetry as the coupling constant varies. It has, however, been successful for only 2 and 3 space-time dimensions.

(2) Analytic S -Matrix Approach. Due to the failure of the perturbation approach in quantum field theory, a new approach was developed based on assumptions about the analyticity properties of the S -matrix elements. The assumed analyticity properties were surmised from examination of Feynman integrals and of nonrelativistic potential scattering, and partially follow from axiomatic field theory. In this approach, the information that scattering amplitudes, or S -matrix elements, possess certain analytic properties with respect to energies, scattering angles, and so on is expressed by means of integral representations. For example, the forward two-body scattering amplitude $f(s)$ as a function of $s=(\text{energy in the center of mass system})^2$ is written as

$$f(s) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\text{Im } f(s')}{s' - s} ds'. \quad (1)$$

$\text{Im } f(s)$ is related to the total cross section by the optical theorem, which is a statement of

the unitarity of the S -matrix. Hence equation (1), called a **dispersion relation**, gives a relation among observable quantities. An integral representation for the general two-body scattering amplitude $f(s, t, u)$ (2 incoming and 2 outgoing particles) has been proposed by S. Mandelstam and is called the **Mandelstam representation**, where $s=(p_1 + p_2)^2$, $t=(p_1 + p_3)^2$, and $u=(p_1 + p_4)^2$ (squares in a Minkowski metric) and p_1, p_2, p_3, p_4 are the 4-momenta of the incoming and outgoing particles, with sign reversed for the latter. (Some relations hold among variables: $p_i^2 = m_i^2$ with m_i the mass, $\sum p_i = 0$, $s + t + u = \sum m_i^2$.) Under the interchange of incoming and outgoing particles, the corresponding f 's are related by analytic continuation. This is called the **crossing symmetry**. The study of the S -matrix directly from its analyticity and unitarity is called the **S -matrix approach** (\rightarrow 386 S -Matrices).

The analyticity of the two-body scattering amplitude f as a function of the angular momentum l with a fixed s was investigated by T. Regge for nonrelativistic potential scattering, and later the idea was applied to the S -matrix approach. The poles $l=l(s)$ of f (f is considered to be a function of l for each fixed s) are called **Regge poles**. Regge trajectories $l=l(s)$ for variable $s \leq 0$ have been shown to play important roles in the high-energy behavior of scattering amplitudes for small values of the variable t (four-momentum-transfer squared).

An approximate expression of S -matrix elements, with Regge poles and satisfying the crossing symmetry, was introduced by G. Veneziano and is called the **Veneziano model**. It has developed into the so-called **dual resonance model** (dual in the sense that s -channel poles are dual to t -channel poles) and has subsequently evolved into the **string model** of hadrons, according to which hadrons are viewed as systems composed of strings joining quarks (and antiquarks).

(3) Group-Theoretical Approach. In connection with the symmetry properties shown by the spectra and reaction patterns of hadrons, a group-theoretical approach has been developed. For example, the similarity of the behavior of neutrons and protons in nuclei, apart from the difference in their electromagnetic properties, was formulated as isospin invariance (under the group $SU(2)$). The symmetry properties are sometimes understood in terms of new additive quantum numbers that are conserved or nearly conserved, and these properties are made more concrete in the final stage by the introduction of fundamental constituents carrying these quantum numbers.

Based on the canonical formalism of quantum field theory, (Noether) currents, as

quantum-mechanical generators of the symmetry, are introduced in association with conserved quantum numbers. The commutation relation of these currents, referred to as a **current algebra**, shows the structure of a Lie algebra expressing the symmetry of the Lagrangian of the system of hadrons; for example, $SU(3) \times SU(3)$ for three species (flavors) of massless quarks (the first factor for vector currents and the second for axial-vector currents). The approach showed remarkable success when combined with the hypothesis of the **partially conserved axial-vector currents** (PCAC), which requires the divergence of the axial-vector currents to be proportional to the pseudoscalar meson fields.

Even if a theory has a symmetry under a group G in its formulation, a vacuum state of the theory might not have a symmetry under G . If that occurs, we speak of **spontaneously broken symmetry**. Under some assumptions, a particle of zero mass (which is connected to the vacuum by the current for the spontaneously broken symmetry) is associated with the spontaneously broken symmetry. This statement is called **Goldstone's theorem**, and the relevant particle of zero mass is called the **Nambu-Goldstone boson**. The PCAC hypothesis is believed to be connected with the spontaneous breakdown of the axial $SU(3)$ symmetry with pions, kaons, and eta-mesons as Nambu-Goldstone bosons. This approach produced the **Adler-Weisberger sum rule**, which relates the weak axial-vector coupling constants to pion-nucleon scattering cross sections.

A group-theoretical attempt to treat fermions and bosons on an equal footing resulted in the introduction of super Lie algebras (Z_2 -graded Lie algebras). The basic new ingredient in this approach is a special class of generators roughly interpreted as the square root of the four-momentum, whose anticommutators (instead of commutators) are linear combinations of ordinary generators. A supermultiplet, which is an irreducible representation of a super Lie algebra, consists of both fermions and bosons. Extensions to local super Lie algebras (and also incorporation of gravitons into the framework) have been tried, with no realistic model emerging so far.

(4) Non-Abelian Gauge Field Theory. Recently, quantum field theory has been revived as non-Abelian gauge theory, resulting in what are considered to be successful qualitative and semiquantitative predictions. The quantization of the theory was at first carried out in terms of the Feynman path integral, where fictitious particles, called **Faddeev-Popov ghosts**, appear through the precise definition of the functional measure of the path integral. The canonical

quantization has also been formulated with the explicit introduction of Faddeev-Popov ghosts from the beginning. Non-Abelian gauge field theory exhibits the very important property of asymptotic freedom, which states that the interaction at asymptotically high energies, or at very short distances, approaches that of free (noninteracting) theory. This property is required for the description of hadronic systems made of quarks in view of the experimental observation of the scaling behavior of deep inelastic inclusive structure functions. Thus gauge theory is believed to describe the dynamics of systems of quarks in hadrons. It is called **quantum chromodynamics**, and the quantum of the gauge field is called the gluon. The recent development of gauge field theory has been accompanied by many technical refinements of quantum field theory, which include the methods of dimensional regularization and renormalization groups.

Dimensional regularization starts with Feynman integrals defined for n -dimensional momenta with $n \neq 4$, in order to give meaning to integrals divergent for $n = 4$. Then Feynman integrals have poles at $n = 4$, which are absorbed into the unrenormalized constants by means of the renormalization procedure. This method is particularly suited for non-Abelian gauge theory, since it is the regularization method that keeps gauge invariance at each step of the calculation.

The renormalization-group equation has been known for a long time. It results from the requirement that physical quantities should be dimensionally covariant under renormalization of constants in the theory. The renormalization-group equation is usually written as a differential equation for a Green's function, expressing the fact that a change of scale of momenta is balanced by changing coupling constants and masses. Further refinements are due to C. G. Callan, K. Symanzik, S. Weinberg, and G. 't Hooft.

D. Models of Elementary Particles

Hadrons are now considered to be made of more fundamental constituents; at present, at least 15 species of subhadronic, fermionic constituents (quarks) have been proposed. Attempts to understand subhadronic structure have a long history, the main landmarks of which include the Fermi-Yang model (where π -mesons are supposed to be made of protons and neutrons), the **Sakata model** (where all hadrons are supposed to be made of protons, neutrons, and Λ -hyperons), and a variety of quark models developed from the **eightfold way** of M. Gell-Mann and Y. Ne'eman (new

assignments of representations of $SU(3)$ to particles somewhat different from the Sakata model; for example, the octet representation for mesons and low-lying baryons, and the decuplet for excited baryons). Originally, quarks were supposed to come in 3 species (u -quarks, d -quarks, s -quarks), each carrying its own quantum number, now called the flavor quantum number. Then it was suggested that each flavored quark has three additional degrees of freedom, i.e., three new quantum numbers (called color quantum numbers) in order to account for experimental data: (1) the spin-statistics problem of baryonic ground state wave functions, (2) the decay rate of $\pi \rightarrow 2\gamma$, (3) the Drell ratio (= total cross section for $\{e^- + e^+ \rightarrow \text{anything}\}$, divided by the cross section for $e^- + e^+ \rightarrow \mu^- + \mu^+$). Recently, two new flavor degrees of freedom besides the old 3 flavors (u, d, s) have been discovered, the carriers of which are c -quarks and b -quarks, c being the constituents of J/ψ -particles, b of Y -particles. The combination of 5 flavors and 3 colors results in 15 quarks, as stated earlier.

The so-called standard model is a quantum field theory based on a local, non-Abelian gauge group $SU(3) \times SU(2) \times U(1)$. The group $SU(3)$ is called the color $SU(3)$ group, which is supposed to be strictly unbroken and expresses the invariance of the theory under the local $SU(3)$ transformation of the three color degrees of freedom. Quarks having spin $\frac{1}{2}$ transform as its 3-dimensional fundamental representation, and the vector gauge bosons transforming as its 8-dimensional regular representation are **gluons**. This part of the theory is **quantum chromodynamics (QCD)**. The remaining part of the theory, based on the local gauge group $SU(2) \times U(1)$, is called the **Glashow-Weinberg-Salam model** or its hadronic extension **quantum flavor dynamics (QFD)**, and unifies the electromagnetic and the weak interactions. This gauge group is supposed to be spontaneously broken with the only unbroken subgroup $U(1)$ corresponding to the electromagnetic gauge transformation. The underlying mechanism for the spontaneous breakdown of the gauge group $SU(2) \times U(1)$ is not well understood, but conventionally it is assumed to occur through the so-called **Higgs mechanism**. This is a mechanism proposed by P. W. Higgs, whereby the Goldstone boson acquires a nonzero mass if the broken symmetry occurs in the presence of an associated massless vector field (called a gauge vector field), which also becomes associated with massive bosons. The gauge vector bosons are identified with photons (γ) (unbroken symmetry and hence massless) and weak intermediary bosons (broken symmetry and hence massive), usually denoted by W and Z . Quarks

and leptons transform under the group $SU(2)$ as doublet or singlet representations. The renormalizability of the spontaneously broken gauge field theory has been established by 't Hooft. Major predictions of the Glashow-Weinberg-Salam model, including the existence of the gauge bosons W and Z , have been borne out experimentally.

Grand unified models attempt to unify QCD and QFD, employing a larger Lie group containing $SU(3) \times SU(2) \times U(1)$ as a subgroup. The most popular ones are those based on the groups $SU(5)$, $SO(10)$ and on some of the exceptional groups. Super grand unified models attempt to unify QCD, QFD, and the gravitational interaction, with super Lie groups as a possible basis. Recently, there have been attempts to search for subquark structures.

References

- [1] J. D. Bjorken and S. D. Drell, *Relativistic quantum mechanics*, McGraw-Hill, 1964.
- [2] J. D. Bjorken and S. D. Drell, *Relativistic quantum fields*, McGraw-Hill, 1965.
- [3] G. F. Chew, *The analytic S-matrix*, Benjamin, 1966.
- [4] S. L. Adler and R. F. Dashen, *Current algebras*, Benjamin, 1967.
- [5] V. de Alfaro, S. Fubini, G. Furlan, and C. Rossetti, *Currents in hadron physics*, North-Holland, 1973.
- [6] M. Gell-Mann and Y. Ne'eman, *The eightfold way*, Benjamin, 1964.
- [7] R. E. Behrends, J. Dreitlein, C. Fronsdal, and W. Lee, Simple groups and strong interaction symmetries, *Rev. Mod. Phys.*, 34 (1962), 1–40.
- [8] E. Fermi and C. N. Yang, Are mesons elementary particles? *Phys. Rev.*, 76 (1949), 1739–1743.
- [9] S. Sakata, On a composite model for the new particles, *Prog. Theoret. Phys.*, 16 (1956), 686–688.
- [10] L. Corwin, Y. Ne'eman, and S. Sternberg, Graded Lie algebra in mathematics and physics, *Rev. Mod. Phys.*, 47 (1975), 573–603.
- [11] M. Jacob (ed.), *Dual theory*, Physics Reports Reprint Book Series, vol. 1, North-Holland, 1974.
- [12] J. Scherk, An introduction to the theory of dual models and strings, *Rev. Mod. Phys.*, 47 (1975), 123–164.
- [13] H. Harari, Quarks and leptons, *Phys. Rep.*, 42 (1978), 235–309.
- [14] E. S. Abers and B. W. Lee, Gauge theories, *Phys. Rep.*, 9 (1973), 1–141.
- [15] S. Weinberg, Recent progress in gauge theories of the weak, electromagnetic and

strong interactions, Rev. Mod. Phys., 46 (1974), 255–277.
[16] J. C. Taylor, Gauge theories of weak interactions, Cambridge Univ. Press, 1976.

133 (XIV.9) Ellipsoidal Harmonics

A. Ellipsoidal Coordinates

If $a > b > c$, then for any given $(x, y, z) \in \mathbf{R}^3$, the three roots of the cubic equation in θ

$$F(\theta) = \frac{x^2}{a^2 + \theta} + \frac{y^2}{b^2 + \theta} + \frac{z^2}{c^2 + \theta} - 1 = 0$$

are real and lie in the intervals $\theta > -c^2$, $-c^2 > \theta > -b^2$, and $-b^2 > \theta > -a^2$. Denoting these three roots by λ , μ , and ν (they are labeled so as to satisfy the inequalities $\lambda > -c^2 > \mu > -b^2 > \nu > -a^2$), $F(\lambda) = 0$, $F(\mu) = 0$, and $F(\nu) = 0$ represent an ellipsoid, a hyperboloid of one sheet, and a hyperboloid of two sheets, respectively. They are confocal with the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0,$$

pass through the point (x, y, z) , and mutually intersect orthogonally.

The quantities λ , μ , ν are called the **ellipsoidal coordinates** of the point (x, y, z) . Rectangular coordinates (x, y, z) are expressed in terms of ellipsoidal coordinates (λ, μ, ν) by the formula

$$x^2 = \frac{(a^2 + \lambda)(a^2 + \mu)(a^2 + \nu)}{(a^2 - b^2)(a^2 - c^2)}, \quad (1)$$

and two others obtained from (1) by cyclic permutations of (a, b, c) and (x, y, z) .

B. Ellipsoidal Harmonics

When a \dagger harmonic function ψ of three real variables is constant on the surface $\lambda = \text{constant}$, $\mu = \text{constant}$, or $\nu = \text{constant}$ in ellipsoidal coordinates, the function ψ is called an **ellipsoidal harmonic**. A solution of Laplace's equation $\Delta\psi = 0$ in the form $\psi = \Lambda(\lambda)M(\mu)N(\nu)$ can be obtained by the method of separation of variables. The equation $\Delta\psi = 0$ is written in the form

$$\sum (\mu - \nu) \Delta_\lambda \frac{\partial}{\partial \lambda} \left(\Delta_\lambda \frac{\partial \psi}{\partial \lambda} \right) = 0,$$

where the summation is taken over the even permutations of (λ, μ, ν) , and

$$\Delta_\lambda = \sqrt{(a^2 + \lambda)(b^2 + \lambda)(c^2 + \lambda)}.$$

The ordinary differential equation

$$4\Delta_\lambda \frac{d}{d\lambda} \left(\Delta_\lambda \frac{d\Lambda}{d\lambda} \right) = (K\lambda + C)\Lambda \quad (2)$$

is satisfied by Λ and also by M and N if we replace λ by μ and ν , respectively. Equation (2) is called **Lamé's differential equation**, with K and C the separation constants.

Let $K = n(n+1)$ for $n = 0, 1, 2, \dots$. Then equation (2), for a suitable value (the eigenvalue) of C , has a solution that is a polynomial in λ or a polynomial multiplied by one, two, or three of $\sqrt{a^2 + \lambda}$, $\sqrt{b^2 + \lambda}$, and $\sqrt{c^2 + \lambda}$.

Among these solutions $2n+1$ are linearly independent. We denote these solutions by $\Lambda = f_n^m(\lambda)$ ($m = 1, 2, \dots, 2n+1$). They are essentially equivalent to the Lamé functions, to be defined at the end of this section. To be precise, by setting

$$\lambda + (a^2 + b^2 + c^2)/3 = \xi,$$

$$C = B + n(n+1)(a^2 + b^2 + c^2)/3,$$

$$e_1 = (b^2 + c^2 - 2a^2)/3, \dots, \quad e_1 + e_2 + e_3 = 0,$$

we have

$$\begin{aligned} \frac{d^2 \Lambda}{d\xi^2} + \left(\frac{1/2}{\xi - e_1} + \frac{1/2}{\xi - e_2} + \frac{1/2}{\xi - e_3} \right) \frac{d\Lambda}{d\xi} \\ = \frac{n(n+1)\xi + B}{4(\xi - e_1)(\xi - e_2)(\xi - e_3)} \Lambda. \end{aligned} \quad (3)$$

This can also be written in the form

$$\frac{d^2 \Lambda}{du^2} = (n(n+1)\wp(u) + B)\Lambda \quad (4)$$

by the change of variable $\xi = \wp(u)$ with the Weierstrass \wp -function.

The differential equation (3) has $\xi = e_1, e_2, e_3, \infty$ as \dagger regular singular points. A solution of (3) that is a polynomial in ξ or a polynomial multiplied by one, two, or three of $\sqrt{\xi - e_1}$, $\sqrt{\xi - e_2}$, and $\sqrt{\xi - e_3}$ is called a **Lamé function of the first kind**.

C. Classification of Lamé Functions

The $2n+1$ linearly independent solutions $f_n^m(\lambda)$ of (2) are classified into the following four families. If n is an even number $2p$, then $p+1$ solutions $f_n^m(\lambda)$ among the $2n+1$ solutions are polynomials in λ of degree p , and the other $3p$ functions are polynomials in λ of degree $p-1$ multiplied by

$$\sqrt{(b^2 + \lambda)(c^2 + \lambda)}, \quad \sqrt{(c^2 + \lambda)(a^2 + \lambda)},$$

or

$$\sqrt{(a^2 + \lambda)(b^2 + \lambda)}.$$

Since all these polynomials are products with real factors of degree 1, the solutions belonging

to the first family are of the type

$$f_n^m(\lambda) = (\lambda - \theta_1)(\lambda - \theta_2) \dots (\lambda - \theta_{n/2}), \quad (5)$$

while solutions of the latter kinds are of the type

$$f_n^m(\lambda) = \left\{ \begin{array}{l} \sqrt{(b^2 + \lambda)(c^2 + \lambda)} \\ \sqrt{(c^2 + \lambda)(a^2 + \lambda)} \\ \sqrt{(a^2 + \lambda)(b^2 + \lambda)} \end{array} \right\} \times (\lambda - \theta_1)(\lambda - \theta_2) \dots (\lambda - \theta_{n/2-1}). \quad (6)$$

The functions (5) and (6) are called **Lamé functions of the first species and of the third species**, respectively. On the other hand, if n is an odd number $2p + 1$, then $3(p + 1)$ solutions among the $2n + 1$ functions $f_n^m(\lambda)$ are of the type

$$f_n^m(\lambda) = \left\{ \begin{array}{l} \sqrt{a^2 + \lambda} \\ \sqrt{b^2 + \lambda} \\ \sqrt{c^2 + \lambda} \end{array} \right\} \times (\lambda - \theta_1)(\lambda - \theta_2) \dots (\lambda - \theta_{(n-1)/2}), \quad (7)$$

and the other p functions are of the type

$$f_n^m(\lambda) = \sqrt{(a^2 + \lambda)(b^2 + \lambda)(c^2 + \lambda)} \times (\lambda - \theta_1)(\lambda - \theta_2) \dots (\lambda - \theta_{(n-3)/2}). \quad (8)$$

The functions (7) and (8) are called **Lamé functions of the second species and of the fourth species**, respectively. Hence in either case we have $2n + 1$ linearly independent Lamé functions.

When n is even, we obtain an ellipsoidal harmonic

$$\psi_n^m = \prod_{p=1}^{n/2} (\lambda - \theta_p)(\mu - \theta_p)(v - \theta_p)$$

by multiplying $f_n^m(\lambda)$, $f_n^m(\mu)$, and $f_n^m(v)$ belonging to the first family. Also, in this case, by setting

$$\Theta_p = \frac{x^2}{a^2 + \theta_p} + \frac{y^2}{b^2 + \theta_p} + \frac{z^2}{c^2 + \theta_p} - 1 \\ = \frac{(\lambda - \theta_p)(\mu - \theta_p)(v - \theta_p)}{(a^2 + \theta_p)(b^2 + \theta_p)(c^2 + \theta_p)}$$

we have

$$\psi_n^m = \Theta_1 \Theta_2 \dots \Theta_{n/2} \quad (9)$$

up to constant coefficients. Utilizing Lamé functions of the third species (instead of functions of the first species) and formula (1), we find that

$$\psi_n^m = (yz \text{ or } zx \text{ or } xy) \times \Theta_1 \Theta_2 \dots \Theta_{n/2-1}. \quad (10)$$

For even n , every ellipsoidal harmonic expressible in terms of polynomials in x , y , z of degree n can be written as a linear combination of the functions (9) and (10), which are

called the **ellipsoidal harmonics of the first and of the third species**, respectively. Similarly for odd n , the **ellipsoidal harmonics of the second and of the fourth species**

$$\psi_n^m = (x \text{ or } y \text{ or } z) \Theta_1 \Theta_2 \dots \Theta_{(n-1)/2}, \quad (11)$$

$$\psi_n^m = xyz \Theta_1 \Theta_2 \dots \Theta_{(n-3)/2} \quad (12)$$

are composed of the Lamé functions of the second and fourth species, respectively.

For odd n , these forms are a complete system of ellipsoidal harmonics that are linearly independent and expressible in terms of polynomials in x , y , z of degree n .

The zeros $\xi_1, \xi_2, \dots, \xi_p$ of the Lamé functions are real, and $\xi_i \neq \xi_j$ ($i \neq j$). They never coincide with any one of e_1, e_2 , and e_3 . If $e_1 > e_2 > e_3$, then ξ_1, \dots, ξ_p all lie between e_1 and e_3 . If m is an integer such that $0 \leq m \leq p$, we have one and only one Lamé function (with the species given) with m of its zeros lying between e_1 and e_2 and the remaining $p - m$ zeros between e_2 and e_3 (**Stieltjes's theorem**). In this way, a complete system of linearly independent Lamé functions of the specified type may be obtained, since m assumes $p + 1$ different values. When the constant B appearing in the differential equation (3) takes specific values so that the equation has Lamé functions of the first kind as its solution, (3) also yields a solution Λ such that $\Lambda \rightarrow \xi^{-(n+1)/2}$ as $\xi \rightarrow \infty$. This function Λ is called the **Lamé function of the second kind**.

D. Ellipsoids of Revolution (Spheroids)

When the fundamental ellipsoid is a spheroid

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = 1,$$

it is convenient to use the **spheroidal coordinates** (ξ, η, φ) given by

$$x = l \sqrt{(\xi^2 - 1)(1 - \eta^2)} \cos \varphi, \\ y = l \sqrt{(\xi^2 - 1)(1 - \eta^2)} \sin \varphi, \\ z = l \xi \eta, \quad l = \sqrt{c^2 - a^2} \quad (13)$$

for $a^2 < c^2$ (prolate) and

$$x = l \sqrt{(\xi^2 + 1)(1 - \eta^2)} \cos \varphi, \\ y = l \sqrt{(\xi^2 + 1)(1 - \eta^2)} \sin \varphi, \\ z = l \xi \eta, \quad l = \sqrt{a^2 - c^2} \quad (14)$$

for $a^2 > c^2$ (oblate). The solutions of Laplace's equation, which are regular at all finite points, are given by

$$\psi = P_n^m(\xi) P_n^m(\eta) \cos m \varphi$$

in the prolate case, and

$$\psi = P_n^m(i\xi) P_n^m(\eta) \cos m \varphi$$

in the oblate case. Here P_n^m is the †associated Legendre function of the first kind. Solutions which are regular outside a finite ellipsoid can be composed of the †associated Legendre functions of the second kind, $Q_n^m(\xi)$ or $Q_n^m(i\xi)$ instead of $P_n^m(\xi)$ or $P_n^m(i\xi)$, respectively.

E. Spheroidal Wave Functions

Transforming the †Helmholtz equation in prolate spheroidal coordinates (13), we have

$$\frac{1}{(\xi^2-\eta^2)}\left\{\frac{\partial}{\partial\xi}\left((\xi^2-1)\frac{\partial\Psi}{\partial\xi}\right)+\frac{\partial}{\partial\eta}\left((1-\eta^2)\frac{\partial\Psi}{\partial\eta}\right)+\left(\frac{1}{\xi^2-1}+\frac{1}{1-\eta^2}\right)\frac{\partial^2\Psi}{\partial\varphi^2}\right\}+\kappa^2\Psi=0,\quad \kappa=kl. \tag{15}$$

By separating variables in the form $\Psi=X(\xi)Y(\eta)\cos m\varphi$, we have the equations

$$\frac{d}{d\xi}\left((1-\xi^2)\frac{dX}{d\xi}\right)+\left(\lambda-\kappa^2\xi^2-\frac{m^2}{1-\xi^2}\right)X=0, \tag{16a}$$

$$\frac{d}{d\eta}\left((1-\eta^2)\frac{dY}{d\eta}\right)+\left(\lambda-\kappa^2\eta^2-\frac{m^2}{1-\eta^2}\right)Y=0, \tag{16b}$$

which X and Y , respectively, must satisfy. The only difference between equations (16a) and (16b) arises from the fact that the domain of (16a) is given by $1<\xi$ whereas the domain of (16b) is given by $-1<\eta<1$. For the oblate spheroid, utilizing formula (14) we have

$$\frac{1}{(\xi^2+\eta^2)}\left\{\frac{\partial}{\partial\xi}\left((\xi^2+1)\frac{\partial\Psi}{\partial\xi}\right)+\frac{\partial}{\partial\eta}\left((1-\eta^2)\frac{\partial\Psi}{\partial\eta}\right)+\left(\frac{1}{1-\eta^2}-\frac{1}{\xi^2+1}\right)\frac{\partial^2\Psi}{\partial\varphi^2}\right\}+\kappa^2\Psi=0. \tag{17}$$

By separating variables as before, $Y(\eta)$ satisfies the same equation as (16b), while $X(\xi)$ satisfies equation (16a) with ξ replaced by $i\xi$. All these equations are of the type

$$\frac{d}{dz}\left((1-z^2)\frac{du}{dz}\right)+\left(\lambda-\kappa^2z^2-\frac{m^2}{1-z^2}\right)u=0. \tag{18}$$

A solution of (18) is known as a **spheroidal wave function**. The equation (18) has ± 1 as regular singular points and ∞ as an irregular singular point of class 1. Hence spheroidal wave functions behave like †Legendre functions in the interval $[-1, 1]$ and like †Bessel functions in the neighborhood of ∞ .

F. The Functions $pe_n^m(x)$ and $qe_n^m(x)$

When we write a solution of (18), it is customary to write x instead of z when z is contained in the interval $[-1, 1]$. We denote solutions of

(18) which are regular on the whole domain $-1\leq x\leq 1$ by $pe_n^m(x)$, and the corresponding eigenvalues by $\lambda_{n,m}$ (assuming the boundary condition stated in this paragraph concerning singularities). In particular, when $\kappa\rightarrow 0$, equation (18) reduces to †Legendre's associated differential equation, the eigenvalues of λ become $n(n+1)$ (n is a positive integer), and the corresponding eigenfunctions become the associated Legendre functions of the first kind:

$$P_n^m(x)=(1-x^2)^{m/2}\frac{d^m P_n}{dx^m}. \tag{19}$$

Hence $pe_n^m(x)$ is a solution which tends to a constant multiple of $P_n^m(x)$ as $\kappa\rightarrow 0$. Using a system of orthogonal functions $P_n^m(x)$, we can expand $pe_n^m(x)$ as

$$pe_n^m(x)=\sum_{l\geq m}A_{n,l}^mP_l^m(x), \tag{20}$$

$|l-n|=\text{even number.}$

The coefficients A satisfy a recurrence formula

$$\begin{aligned} &\left(\lambda_{n,m}-l(l+1)+\kappa^2\frac{2l^2+2l-1-2m^2}{(2l-1)(2l+3)}\right)A_{n,l}^m \\ &- \kappa^2\frac{(l-m-1)(l-m)}{(2l-3)(2l-1)}A_{n,l-2}^m \\ &- \kappa^2\frac{(l+m+1)(l+m+2)}{(2l+3)(2l+5)}A_{n,l+2}^m=0. \end{aligned} \tag{21}$$

The functions $pe_n^m(x)$ and $pe_l^m(x)$ are orthogonal in the domain $-1< x < 1$.

Another solution of (18) exists which corresponds to the same eigenvalue $\lambda_{n,m}$, is independent of $pe_n^m(x)$, and has the opposite parity:

$$\begin{aligned} qe_n^m(x) &= \sum_{l\geq -m, l-n=\text{even}}A_{n,l}^mQ_l^m(x) \\ &+ \sum_{j\geq m, j-n=\text{odd}}B_{n,j}^mP_j^m(x), \end{aligned} \tag{22}$$

where the $A_{n,l}^m$ are the same as in (20) and are determined by the recurrence formula (21), while for $j\geq m+2$, the $B_{n,j}^m$ satisfy the recurrence formula

$$\begin{aligned} &\left(\lambda_{n,m}-j(j+1)+\kappa^2\frac{2j^2+2j-1-2m^2}{(2j-1)(2j+3)}\right)B_{n,j}^m \\ &+ \kappa^2\frac{(j-m-1)(j-m)}{(2j-3)(2j-1)}B_{n,j-2}^m \\ &+ \kappa^2\frac{(j+m+1)(j+m+2)}{(2j+3)(2j+5)}B_{n,j+2}^m=0. \end{aligned} \tag{23}$$

Since the associated Legendre function of the second kind

$$Q_l^m(x)=(1-x^2)^{m/2}\frac{d^m Q_l}{dx^m}$$

is of the form

$$Q_l^m(x)=P_l^m(x)\log\sqrt{(1+x)/(1-x)}+(1-x^2)^{-m/2}\times(\text{a polynomial in } x)$$

for $l \geq m$, the $q e_n^m(x)$ have $x = \pm 1$ as singular points.

By expressing the solution of equation (18) in integral form we find that $p e_n^m(x)$ satisfies an integral equation

$$i^{n-m} v_{n,m} p e_n^m(x) = \int_{-1}^1 (1-x^2)^{m/2} (1-\xi^2)^{m/2} e^{ikx\xi} p e_n^m(\xi) d\xi, \quad (24)$$

where the coefficient $v_{n,m}$ is related to $P_n^m(0)$ or $P_n^{m'}(0)$.

In order to extend the domain of definition of $p e_n^m(x)$ and $q e_n^m(x)$ beyond the interval $[-1, 1]$, we adopt, in the domain G obtained by deleting the interval $[-1, 1]$ from the complex plane, the Heine-Hobson definition of the associated Legendre function,

$$P_n^m(z) = (z^2 - 1)^{m/2} d^m P_n / dz^m, \quad (25)$$

instead of N. M. Ferrers's definition (19), and construct a solution of (18) in G :

$$p e_n^m(z) = \sum_{l \geq m} A_{n,l}^m P_l^m(z),$$

$$|l - n| = \text{even number}, \quad (26)$$

which is like (20) and again satisfies the integral equation (24). From this we can obtain the expansion formula

$$p e_n^m(z) = \frac{\sqrt{2\pi} (z^2 - 1)^{m/2}}{v_{n,m} (\kappa z)^m} \times \sum_{l \geq m} (-1)^{(l-n)/2} \frac{(l+m)!}{(l-m)!} A_{n,l}^m \frac{J_{l+1/2}(\kappa z)}{\kappa z},$$

$$|l - n| = \text{even number}. \quad (27)$$

Multiplying this by a constant, we define

$$j e_n^m(z) = \sqrt{\frac{\pi}{2}} \frac{(z^2 - 1)^{m/2}}{z^m} \times \sum_{l \geq m} (-1)^{(l-n)/2} F_{n,l}^m \frac{J_{l+1/2}(\kappa z)}{\sqrt{\kappa z}} \bigg/ \sum_{l \geq m} F_{n,l}^m,$$

$$F_{n,l}^m = \frac{(l+m)!}{(l-m)!} A_{n,l}^m.$$

This expression asymptotically assumes the form

$$j e_n^m(z) \sim \sin(\kappa z - n\pi/2) / \kappa z$$

for $|z| \gg 1$. In a similar manner we find a solution

$$n e_n^m(z) = -\sqrt{\frac{\pi}{2}} \frac{(z^2 - 1)^{m/2}}{z^m} \times \sum_{l \geq m} (-1)^{(l-n)/2} F_{n,l}^m \frac{N_{l+1/2}(\kappa z)}{\sqrt{\kappa z}} \bigg/ \sum_{l \geq m} F_{n,l}^m$$

having the asymptotic form

$$n e_n^m(z) \sim \cos(\kappa z - n\pi/2) / \kappa z.$$

Disregarding a constant factor, this coincides with the function defined by (22) with $Q_n^m(z) = (z^2 - 1)^{m/2} d^m Q_n / dz^m$ in place of the associated Legendre function of the second kind.

References

- [1] M. J. O. Strutt, Lamésche- Mathieusche- und verwandte Funktionen in Physik and Technik, Erg. Math., Springer, 1932 (Chelsea, 1967).
- [2] E. W. Hobson, The theory of spherical and ellipsoidal harmonics, Cambridge Univ. Press, 1931 (Chelsea, 1955).
- [3] Y. Hagihara, Theories of equilibrium figures of a rotating homogeneous fluid mass, revised translation, NASA, Washington D.C., 1970. (Original in Japanese, 1933.)
- [4] Y. Hagihara, Celestial mechanics, vol. 4, pts. 1, 2, Japan Society for the Promotion of Science, 1975.
- [5] C. Flammer, Spheroidal wave functions, Stanford Univ. Press, 1957.
- [6] J. A. Stratton, P. M. Morse, L. J. Chu, J. D. C. Little, and F. J. Corbató, Spheroidal wave functions, including tables of separation constants and coefficients, Wiley, 1956.

134 (XIV.3) Elliptic Functions

A. Elliptic Integrals

Let $\varphi(z)$ be a polynomial in z of degree 3 or 4 with complex coefficients and $R(z, w)$ a rational function in z and w . Then $R(z, \sqrt{\varphi(z)})$ is called an **elliptic irrational function**. An integral of the type $\int R dz$ is called an **elliptic integral**. The origin of the name comes from the integral that appears in calculating the arc length of an ellipse. Any elliptic integral can be expressed by a suitable change of variables as a sum of elementary functions and elliptic integrals of the following three kinds:

$$\int \frac{dz}{\sqrt{(1-z^2)(1-k^2 z^2)}},$$

$$\int \sqrt{\frac{1-k^2 z^2}{1-z^2}} dz,$$

and

$$\int \frac{dz}{(1-a^2 z^2) \sqrt{(1-z^2)(1-k^2 z^2)}}$$

(\rightarrow Appendix A, Table 16.I). These three kinds of integrals are called **elliptic integrals of the first, second, and third kind**, respectively, in **Legendre-Jacobi standard form**. This classifica-

tion corresponds to that of \dagger Abelian integrals. The constant k is called the **modulus** of these elliptic integrals, and a is called the **parameter**. Let the four zeros of $\varphi(z)$ be $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ (we take one of them as ∞ when $\varphi(z)$ is of degree 3). The \dagger Riemann surface \mathfrak{R} corresponding to the elliptic irrational function has the zeros $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ as \dagger branch points with degree of ramification 1, and is of two sheets and of \dagger genus 1. If the integrand does not have a pole with a \dagger residue, then the integral is multivalued only because the value (called the **periodicity modulus**) of the integral taken along the normal section (basis of the homology group) is not equal to zero.

B. Elliptic Integrals of the First Kind

When R is a function without singularities other than branch points, only the topological structure of R gives rise to the multivaluedness of the integral of R . The standard form is

$$\begin{aligned} \int_0^z \frac{dz}{\sqrt{(1-z^2)(1-k^2z^2)}} \\ = \int_0^\varphi \frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}} = w = F(k, \varphi), \end{aligned} \quad (1)$$

where $z = \sin \varphi$. This integral is the inverse function of $\operatorname{sn} w$ (\rightarrow Section J). The periodicity moduli are $2iK'$ and $4K$, where

$$\begin{aligned} K = K(k) &= \int_0^1 \frac{dz}{\sqrt{(1-z^2)(1-k^2z^2)}} \\ &= \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}} = F\left(k, \frac{\pi}{2}\right), \end{aligned}$$

$$K' = K(k'), \quad k'^2 = 1 - k^2.$$

We call $K(k)$ a **complete elliptic integral of the first kind** and $F(k, \varphi)$ an **incomplete elliptic integral of the first kind** (\rightarrow Appendix A, Table 16). Setting

$$\sin \varphi_1 = \frac{(1+k')\sin \varphi \cos \varphi}{\sqrt{1-k^2\sin^2\varphi}}, \quad k_1 = \frac{1-k'}{1+k'}, \quad (2)$$

we have the relation

$$F(k, \varphi) = (1+k_1)F(k_1, \varphi_1)/2,$$

which is called **Landen's transformation**. Since $k_1 < k$ when $0 < k < 1$, this transformation reduces the calculation of elliptic integrals to those with smaller values of k . For two given positive numbers a and b , put $a_0 = a$, $b_0 = b$, and $a_{n+1} = (a_n + b_n)/2$, $b_{n+1} = \sqrt{a_n b_n}$. Then the sequences $\{a_n\}$ and $\{b_n\}$ converge rapidly to the common limit, which is called the **arithmetic-geometric mean** of a and b , and is denoted by $\operatorname{ag}(a, b)$. The complete

elliptic integral satisfies the relation $K(k) = \pi/[2\operatorname{ag}(1, \sqrt{1-k^2})]$.

C. Elliptic Integrals of the Second Kind

When R has poles with residue zero, its integral has no singularities other than poles.

The standard form is

$$\begin{aligned} F(z) &= \int_0^z \sqrt{\frac{1-k^2z^2}{1-z^2}} dz \\ &= \int_0^\varphi \sqrt{1-k^2\sin^2\varphi} d\varphi = E(k, \varphi), \end{aligned} \quad (3)$$

where $z = \sin \varphi$. We have

$$F(z) = \int_0^u \operatorname{dn}^2 u du = \frac{\Theta'(u)}{\Theta(u)} + \frac{E}{K} u$$

if we set $z = \operatorname{sn} u$. Here, Θ is Jacobi's theta function, and

$$\Theta(u) = \mathfrak{g}_4\left(\frac{u}{2K}, \frac{iK'}{K}\right),$$

where \mathfrak{g}_4 is a theta function to be described in Section I and K, K' are the same as in the case of elliptic integrals of the first kind. The quantity

$$\begin{aligned} E &= \int_0^1 \sqrt{\frac{1-k^2z^2}{1-z^2}} dz \\ &= \int_0^{\pi/2} \sqrt{1-k^2\sin^2\varphi} d\varphi = E\left(k, \frac{\pi}{2}\right) \end{aligned}$$

is called a **complete elliptic integral of the second kind**.

D. Elliptic Integrals of the Third Kind

When R has poles with nonzero residue its integral has logarithmic singularities. In this case, residues also contribute to multivaluedness of the integral. The standard form is

$$\begin{aligned} F(z) &= \int_0^z \frac{dz}{(1-a^2z^2)\sqrt{(1-z^2)(1-k^2z^2)}} \\ &= \int_0^\varphi \frac{d\varphi}{(1-a^2\sin^2\varphi)\sqrt{1-k^2\sin^2\varphi}}, \end{aligned}$$

and it is expressed as

$$F(z) = \frac{\operatorname{sn} \alpha}{\operatorname{cn} \alpha \operatorname{dn} \alpha} \left(\frac{1}{2} \log \frac{\Theta(u-\alpha)}{\Theta(u+\alpha)} + u \frac{\Theta'(u)}{\Theta(u)} \right) + u$$

if we set $z = \operatorname{sn} u$ and $a^2 = k^2 \operatorname{sn}^2 \alpha$ (\rightarrow Appendix A, Table 16).

E. Elliptic Functions and Periodic Functions

Historically the elliptic function was first introduced as the inverse function of the elliptic

integral. However, since it has been realized that elliptic functions are characterized as functions with double periodicity, it is now customary to define them as doubly periodic functions.

If $f(x)$, defined on a linear space X , satisfies the relation $f(x + \omega) = f(x)$ for some $\omega \in X$ and all $x \in X$, the number ω is called a **period** of $f(x)$, and $f(x)$ with a period other than zero is called a **periodic function**. The set P of all periods of $f(x)$ forms an additive group contained in X . If a basis $\omega_1, \dots, \omega_n$ of the additive group P exists, its members are called **fundamental periods** of $f(x)$.

Any continuous, nonconstant, periodic function of a real variable has only one positive fundamental period and is called a **simply periodic function**. The trigonometric functions are typical examples: $\sin x$ and $\cos x$ have the fundamental period 2π ; $\tan x$ and $\cot x$ have the fundamental period π (\rightarrow 159 Fourier Series).

A single-valued nonconstant meromorphic function of n complex variables cannot have more than $2n$ fundamental periods that are linearly independent on the real number field. A function of one complex variable with two fundamental periods is called a **doubly periodic function**.

Let ω, ω' be the fundamental periods of a doubly periodic function. For a given number a , the parallelogram with vertices $a, a + \omega, a + \omega', a + \omega + \omega'$ is called the **fundamental period parallelogram**. The complex plane is covered with a network of congruent parallelograms, called **period parallelograms**, obtained by translating the fundamental period parallelogram through $m\omega + n\omega'$ ($m, n = 0, \pm 1, \pm 2, \dots$).

A doubly periodic function $f(u)$ meromorphic on the complex plane is called an **elliptic function**. For simplicity, we usually denote the fundamental periods of an elliptic function by $2\omega_1$ and $2\omega_3$, and introduce ω_2 defined by the relation $\omega_1 + \omega_2 + \omega_3 = 0$. The first, and therefore also higher, derivatives of any elliptic function are elliptic functions with the same periods. The set of all elliptic functions with the same periods forms a field. The number of poles in a period parallelogram is finite. The sum of the orders of the poles is called the **order** of the elliptic function. An elliptic function with no poles in a period parallelogram is merely a constant (**Liouville's first theorem**). The sum of the residues of an elliptic function at its poles in any period parallelogram is zero (**Liouville's second theorem**). Hence there can be no elliptic function of order 1. An elliptic function of order n assumes any value n times in a period parallelogram (**Liouville's third theorem**). The sum of the zeros minus the

sum of the poles is a period (**Liouville's fourth theorem**).

F. Weierstrass's Elliptic Functions

Weierstrass defined

$$\wp(u) = \frac{1}{u^2} + \sum' \left(\frac{1}{(u - \Omega)^2} - \frac{1}{\Omega^2} \right)$$

as the simplest kind of elliptic function. Here $\Omega = 2m\omega_1 + 2n\omega_3$, with m, n integers. The summation \sum' extends over all integral values (positive, negative, and zero) of m and n , except for $m = n = 0$. $\wp(u)$ is an elliptic function of order 2 with periods $2\omega_1$ and $2\omega_3$, called a **Weierstrass \wp -function**. The following functions $\zeta(u)$ and $\sigma(u)$ are called the **Weierstrass zeta** and **sigma functions**, respectively:

$$\zeta(u) = \frac{1}{u} + \sum' \left(\frac{1}{u - \Omega} + \frac{u}{\Omega^2} + \frac{1}{\Omega} \right)$$

and

$$\sigma(u) = u \prod' \left(\left(1 - \frac{u}{\Omega} \right) \exp \left(\frac{u}{\Omega} + \frac{u^2}{\Omega^2} \right) \right).$$

These have quasiperiodicity, expressed by

$$\zeta(u + 2\omega_i) = \zeta(u) + 2\eta_i, \quad (4)$$

$$\sigma(u + 2\omega_i) = -e^{2\eta_i(u + \omega_i)} \sigma(u), \quad (5)$$

$$\eta_1 + \eta_2 + \eta_3 = 0, \quad \eta_i = \zeta(\omega_i), \quad i = 1, 2, 3;$$

and they satisfy the relations

$$\wp(u) = -\zeta'(u) \quad (6)$$

and

$$\zeta(u) = \frac{d \log \sigma(u)}{du} = \frac{\sigma'(u)}{\sigma(u)}. \quad (7)$$

The function $\wp(u)$ is an even function of u , and $\zeta(u)$ and $\sigma(u)$ are odd functions of u . By considering the integral $\int \zeta(u) du$ once around the boundary of a fundamental period parallelogram, we have

$$\left. \begin{aligned} \eta_1 \omega_3 - \eta_3 \omega_1 \\ \eta_2 \omega_1 - \eta_1 \omega_2 \\ \eta_3 \omega_2 - \eta_2 \omega_3 \end{aligned} \right\} = \pm \frac{\pi}{2} i, \quad \text{Im} \left(\frac{\omega_3}{\omega_1} \right) \geq 0, \quad (8)$$

which is called the **Legendre relation**.

The derivative

$$\wp'(u) = -2 \sum 1/(u - \Omega)^3$$

of a \wp -function is an elliptic function of order 3 and bears the following relation to $\wp(u)$:

$$\begin{aligned} (\wp'(u))^2 &= 4(\wp(u))^3 - g_2 \wp(u) - g_3 \\ &= 4(\wp(u) - e_1)(\wp(u) - e_2)(\wp(u) - e_3), \end{aligned}$$

$$g_2 = 60 \sum' 1/\Omega^4, \quad g_3 = 140 \sum' 1/\Omega^6,$$

$$e_i = \wp(\omega_i), \quad i = 1, 2, 3. \quad (9)$$

By differentiating this relation successively, we see that $\wp^{(n)}(u)$ is expressed as a polynomial in $\wp(u)$ if n is an even number and as a product of polynomials in $\wp(u)$ and $\wp'(u)$ if n is an odd number.

In particular, writing $\wp(u) = z$ in (9) we find that the \wp -function is the inverse function of the elliptic integral

$$u = \int_{\infty}^z \frac{dz}{\sqrt{4z^3 - g_2z - g_3}}$$

(→ Appendix A, Table 16.IV).

Any elliptic function can be expressed in terms of Weierstrass's functions. Specifically, let the poles of $f(u)$ and their orders be a_1, a_2, \dots, a_m and h_1, h_2, \dots, h_m , respectively, and let the principal part in the expansion of $f(u)$ near the pole a_k be

$$\sum_{j=1}^{h_k} \frac{A_{kj}}{(u-a_k)^j}, \quad k = 1, 2, \dots, m. \quad (10)$$

Then we obtain

$$f(u) = C + \sum_{k=1}^m \left(A_{k1} \zeta(u-a_k) + \sum_{j=1}^{h_k} \frac{(-1)^j A_{kj}}{(j-1)!} \wp^{(j-2)}(u-a_k) \right), \quad (11)$$

where C is a constant depending on $f(u)$. This can be reduced to

$$f(u) = A + B\wp'(u)$$

by using the addition theorems (→ Appendix A, Table 16) for \wp and zeta functions, where A and B are rational functions of $D(u)$. Therefore, given two elliptic functions with the same periods, after expressing them as rational functions of \wp and \wp' in the above form and eliminating \wp and \wp' , we obtain an algebraic equation with constant coefficients. In particular, for any elliptic function $f(u)$ we obtain an algebraic differential equation of the first order by using this method, with $f'(u)$ an elliptic function with the same periods. Furthermore, the functions $f(u+v)$, $f(u)$, and $f(v)$ satisfy an algebraic equation. Thus for any elliptic function an algebraic addition theorem holds.

G. Elliptic Functions of the Second Kind

As an extension of the definition of elliptic functions, if a meromorphic function f satisfies the relations

$$f(u+2\omega_1) = \mu_1 f(u), \quad f(u+2\omega_3) = \mu_3 f(u) \quad (12)$$

(μ_1 and μ_3 are constants) with the fundamental periods $2\omega_1, 2\omega_3$, we call $f(u)$ an **elliptic function of the second kind**. What we have called

simply an elliptic function can now be called an **elliptic function of the first kind**. For constants ρ and v , the function

$$f(u) = e^{\rho u} \sigma(u-v) / \sigma(u) \quad (13)$$

is an example of an elliptic function of the second kind. In this case

$$\mu_i = e^{2\rho\omega_i - 2v\eta_i}, \quad i = 1, 2. \quad (14)$$

Furthermore, for given constants μ_1 and μ_3 , any elliptic function of the second kind is expressed as the product of an elliptic function of the first kind and the function (13) with ρ and v determined by (14).

H. Elliptic Functions of the Third Kind

If a meromorphic function f satisfies

$$f(u+2\omega_i) = e^{a_i u + b_i} f(u), \quad i = 1, 3 \quad (15)$$

(a_i and b_i are constants) with periods $2\omega_1, 2\omega_3$, we call it an **elliptic function of the third kind** (→ 3 Abelian Varieties I).

The Weierstrass sigma function $\sigma(u)$ is an example of an elliptic function of the third kind. The functions σ_1, σ_2 , and σ_3 , defined by the equations

$$\sigma_i(u) = -\frac{e^{\eta_i u} \sigma(u-\omega_i)}{\sigma(\omega_i)}, \quad i = 1, 2, 3, \quad (16)$$

are also elliptic functions of the third kind. In the case of elliptic functions of the second and third kinds, $2\omega_1$ and $2\omega_3$ are not periods in the strict sense defined earlier, but are conveniently referred to as the periods. The functions σ_i ($i = 1, 2, 3$) are called **cosigma functions**.

I. Theta Functions

The **theta functions**, or more strictly, **elliptic theta functions**, are defined by

$$\vartheta_1(v, \tau) = 2 \sum_{n=0}^{\infty} (-1)^n q^{(n+1/2)^2} \sin(2n+1)\pi v,$$

$$\vartheta_2(v, \tau) = 2 \sum_{n=0}^{\infty} q^{(n+1/2)^2} \cos(2n+1)\pi v,$$

$$\vartheta_3(v, \tau) = 1 + 2 \sum_{n=1}^{\infty} q^{n^2} \cos 2n\pi v,$$

$$\vartheta_4(v, \tau) = 1 + 2 \sum_{n=1}^{\infty} (-1)^n q^{n^2} \cos 2n\pi v, \quad (17)$$

where $q = e^{i\pi\tau}$, $\text{Im } \tau > 0$. We call (17) the **q -expansion formulas** of the theta functions, and we sometimes write ϑ_0 in place of ϑ_4 . A theta function is an elliptic function of the third kind with periods 1 and τ . Any elliptic function can be expressed as a quotient of theta functions (→ Appendix A, Table 16 for specific exam-

ples). The q -expansion formula is quite suitable for numerical computation because of its rapid convergence; its terms decrease as the n^2 powers of q as $n \rightarrow \infty$.

An elliptic function with the fundamental periods $2\omega_1$ and $2\omega_3$ can also be viewed as having the periods $2\omega'_1 = 2\omega_3$, $2\omega'_3 = -2\omega_1$. Consequently, theta functions formed with the parameter $\tau = \omega_3/\omega_1$ can be expressed in terms of the parameter $\tau' = \omega'_3/\omega'_1 = -\omega_1/\omega_3 = -1/\tau$, and we have

$$\begin{aligned}\vartheta_1(v, \tau) &= iA\vartheta_1(v/\tau, -1/\tau), \\ \vartheta_2(v, \tau) &= A\vartheta_4(v/\tau, -1/\tau), \\ \vartheta_3(v, \tau) &= A\vartheta_3(v/\tau, -1/\tau), \\ \vartheta_4(v, \tau) &= A\vartheta_2(v/\tau, -1/\tau), \\ A &= \sqrt{i/\tau} \exp(-\pi i v^2/\tau).\end{aligned}\quad (18)$$

These are called **Jacobi's imaginary transformations**. If $\text{Im}(-1/\tau) \gg 1$, then $|q| \approx 1$, and therefore the series in the q -expansion formula converges slowly. However, by an imaginary transformation we get $\text{Im}(\tau) \gg 1$ and $|q| \approx 0$, so that computations become much easier.

Each of the theta functions satisfies, as a function of two variables u and τ , the following partial differential equation of the heat-conduction type:

$$\partial^2 \vartheta(u, \tau) / \partial u^2 = 4\pi i \partial \vartheta(u, \tau) / \partial \tau \quad (19)$$

(\rightarrow Appendix A, Table 16.II).

J. Jacobi's Elliptic Functions

C. G. J. Jacobi defined elliptic integrals as inverse functions of elliptic integrals of the first kind in the Legendre-Jacobi standard form (1). They are, in the above notation,

$$\text{sn } w = \sqrt{e_1 - e_3} \frac{\sigma(u)}{\sigma_3(u)} = \frac{\vartheta_3(0)\vartheta_1(v)}{\vartheta_2(0)\vartheta_4(v)}, \quad (20)$$

$$\text{cn } w = \frac{\sigma_1(u)}{\sigma_3(u)} = \frac{\vartheta_4(0)\vartheta_2(v)}{\vartheta_2(0)\vartheta_4(v)}, \quad (21)$$

$$\text{dn } w = \frac{\sigma_2(u)}{\sigma_3(u)} = \frac{\vartheta_4(0)\vartheta_3(v)}{\vartheta_3(0)\vartheta_4(v)}, \quad (22)$$

where $w = \sqrt{e_1 - e_3} u$ and $v = u/2\omega_1$.

These functions satisfy the relations

$$\text{sn}^2 w + \text{cn}^2 w = 1, \quad k^2 \text{sn}^2 w + \text{dn}^2 w = 1, \quad (23)$$

where

$$k^2 = \frac{e_2 - e_3}{e_1 - e_3} = \frac{(\vartheta_2(0))^4}{(\vartheta_4(0))^4}. \quad (24)$$

The constants k and $k' = \sqrt{1 - k^2}$ are called the **modulus** and the **complementary mod-**

ulus, respectively. Furthermore, the relation $d \text{sn } w / dw = \text{cn } w \text{dn } w$ holds. The function $z = \text{sn } w$ is the inverse function of the elliptic integral (1) (\rightarrow Appendix A, Table 16.III).

References

- [1] S. Lang, *Elliptic functions*, Addison-Wesley, 1973.
- [2] D. W. Masser, *Elliptic functions and transcendence*, Springer, 1975.
- [3] C. G. J. Jacobi, *Fundamenta nova theoriae functionum ellipticarum* (1829) (C. G. J. Jacobis Gesammelte Werke I, G. Reimer, 1881, 49–239).
- [4] G. H. Halphen, *Traité des fonctions elliptiques et de leurs applications* I, II, III, Gauthier-Villars, 1886–1891.
- [5] A. Hurwitz and R. Courant, *Vorlesungen über allgemeine Funktionentheorie und elliptische Funktionen*, Springer, third edition, 1929.
- [6] H. E. Rauch and A. Lebowitz, *Elliptic functions, theta functions and Riemann surfaces*, Williams & Wilkins, 1973.
- [7] H. Hancock, *Elliptic integral*, Dover, 1958.

135 (II.3) Equivalence Relations

A. General Remarks

Suppose that we are given a relation R between elements of a set X such that for any elements x and y of X , either xRy or its negation holds. The relation R is called an **equivalence relation** (on X) if it satisfies the following three conditions: (1) xRx , (2) xRy implies yRx , and (3) xRy and yRz imply xRz . Conditions (1), (2), and (3) are called the **reflexive**, **symmetric**, and **transitive laws**, respectively. Together, they are called the **equivalence properties**. Condition (1) can be replaced by the following: (1') For each x there exists an x' such that xRx' . The relation " x is equal to y " is an equivalence relation. If xRy means that x and y are in X , then R is also an equivalence relation. An equivalence relation is often denoted by the symbol \sim . The relations of congruence and similarity between figures are equivalence relations. If X is the set of integers and $x \equiv y$ means that $x - y$ is even, then the relation \equiv is an equivalence relation.

B. Equivalence Classes and Quotient Sets

Let R be an equivalence relation. “ xRy ” is read: “ x and y are **equivalent**” (or “ x is **equivalent to** y ”). The subset of X consisting of all elements equivalent to an element a is called the **equivalence class** of a . By (1), (2), and (3), each equivalence class is nonempty, the equivalence class of a contains a , and different equivalence classes do not overlap. Namely, X is decomposed into a disjoint union of equivalence classes. This partition is called the **classification** of X with respect to R . For example, the set of integers is classified into the equivalence class of even numbers and that of odd numbers by the relation \equiv . Conversely, since the relation “ x and y belong to the same member of a partition” is an equivalence relation, we can regard any partition as a classification. An element chosen from an equivalence class is called a **representative** of the equivalence class. In the example we can take 0 and 1 as the representatives of equivalence classes of even and odd numbers, respectively.

X/R denotes the set of equivalence classes of X with respect to R , and is called the **quotient set** of X with respect to R . The mapping $p: X \rightarrow X/R$ that carries x in X into the equivalence class of x is called **canonical surjection** (or **projection**). The idea of equivalence relations can be generalized to deal with the case when X is a \dagger class.

C. Stronger and Weaker Equivalence Relations

Let R and S be two equivalence relations on X . If xRy always implies xSy , then we say that R is **stronger** than S , S is **weaker** than R , the classification with respect to R is **finer** than the one with respect to S , or the classification with respect to S is **coarser** than the one with respect to R . The relations “ x is equal to y ” and “ x and y are in X ” are the strongest and the weakest equivalence relations, respectively. Any two equivalence relations on X are ordered by their strength, and the set of equivalence relations on X forms a \dagger complete lattice with respect to this ordering.

References

- [1] N. Bourbaki, *Eléments de mathématique I, Théorie des ensembles*, ch. 2, Actualités Sci. Ind., 1212c, Hermann, second edition, 1960; English translation, *Theory of sets*, Addison-Wesley, 1968.
Also \rightarrow references to 381 Sets.

136 (XVII.15) Ergodic Theory

A. General Remarks

The origin of ergodic theory was the so-called **ergodic hypothesis**, which provided the foundation for classical statistical mechanics as created by L. Boltzmann and J. Gibbs toward the end of the 19th century (\rightarrow 402 Statistical Mechanics). Attempts by various mathematicians to give a rigorous proof of the hypothesis resulted in the **recurrence theorem** of H. Poincaré and C. Carathéodory and the **ergodic theorems** of G. D. Birkhoff and J. von Neumann, which marked the beginnings of ergodic theory as we know it today. As the theory developed it acquired close relationships with other branches of mathematics, for example, the theory of dynamical systems, probability theory, functional analysis, number theory, differential topology, and differential geometry.

The principal object of modern ergodic theory is to study properties of \dagger measurable transformations, particularly transformations with an invariant measure. In most cases, the transformations studied are defined on a **Lebesgue measure space with a finite** (or σ -finite) **measure**. A Lebesgue measure space with a finite measure (σ -finite measure) is a \dagger measure space that is measure-theoretically isomorphic to a bounded interval (to the real line) with the usual \dagger Lebesgue measure, possibly together with an at most countable number of atoms. It is known that any separable complete \dagger metric space with a complete regular Borel \dagger probability measure is a Lebesgue measure space with a finite measure. We assume, unless stated otherwise, that the measure space (X, \mathcal{B}, m) is a Lebesgue measure space. All the subsets of X mentioned are assumed measurable, and a pair of sets or functions that coincide almost everywhere are identified. We use the abbreviation “a.e.” to denote “almost everywhere.”

B. Ergodic Theorems

Let (X, \mathcal{B}, m) be a σ -finite measure space. A transformation φ defined on X is called **measurable** if for every $B \in \mathcal{B}$, $\varphi^{-1}(B) \in \mathcal{B}$. A \dagger bijective transformation φ on X is called **bimeasurable** if both φ and φ^{-1} are measurable. A measurable transformation φ is called **measure-preserving** (or equivalently, the measure m is **invariant** under φ) if $m(\varphi^{-1}(B)) = m(B)$ holds for every B . It is called **nonsingular** if $m(\varphi^{-1}(B)) = 0$ whenever $m(B) = 0$, and **ergodic**

if $m((\varphi^{-1}(B) \cup B) - (\varphi^{-1}(B) \cap B)) = 0$ implies either $m(B) = 0$ or $m(X - B) = 0$.

The **mean ergodic theorem** of von Neumann (*Proc. Nat. Acad. Sci. US*, 18 (1932)) states that if φ is a measure-preserving transformation on (X, \mathcal{B}, m) , then for every function f belonging to the \dagger Hilbert space $L_2(X) = L_2(X, \mathcal{B}, m)$ (\rightarrow 168 Function Spaces), the sequence

$$A_n f(x) = \frac{1}{n} \left(\sum_{k=0}^{n-1} f(\varphi^k x) \right)$$

converges in the \dagger norm of $L_2(X)$ as $n \rightarrow \infty$ to a function f^* that satisfies $f^*(\varphi x) = f^*(x)$ a.e. The **individual (or pointwise) ergodic theorem** of Birkhoff (*Proc. Nat. Acad. Sci. US*, 17 (1931)) states that for every f belonging to the \dagger function space $L_1(X)$, the sequence $A_n f(x)$ converges a.e. to f^* . From either of these theorems it follows that for any set E satisfying $\varphi^{-1}(E) = E$ and $m(E) < \infty$, the limit function f^* satisfies $\int_E f^* dm = \int_E f dm$. In particular, if $m(X) = 1$ and φ is ergodic, then the limit f^* equals the constant $\int f dm$ a.e. This fact therefore gives a mathematical justification to the ergodic hypothesis, which states that the "time mean" $(\sum_{k=0}^{n-1} f(\varphi^k x))/n$ of what is observable over a sufficiently long time can be replaced by the "phase mean" $\int f dm$. Both von Neumann's theorem and Birkhoff's theorem were subsequently generalized in various directions by many authors.

(1) Mean ergodic theorems are concerned with the \dagger strong convergence of the sequence of averages $A_n = (\sum_{k=0}^{n-1} T^k)/n$ of the iterates of a \dagger bounded linear operator T on some \dagger Banach space. A generalization of von Neumann's theorem due to F. Riesz, K. Yosida, and S. Kakutani dispenses with the assumptions that the linear operator T is induced by a measure-preserving transformation φ and that T acts on the Hilbert space $L_2(X)$. A version of this generalization states that if a linear operator T defined on a Banach space \mathcal{X} satisfies the conditions

$$(i) \sup_{n \geq 1} \left\| \frac{1}{n} \sum_{k=0}^{n-1} T^k \right\| < \infty,$$

$$(ii) \lim_{n \rightarrow \infty} \frac{1}{n} \|T^n\| = 0,$$

then for an element $f \in \mathcal{X}$ the sequence of averages $A_n f$ converges strongly to an element $f^* \in \mathcal{X}$ if and only if there exists a subsequence converging \dagger weakly to f^* . From this theorem of Riesz, Yosida, and Kakutani follows the L_p -mean ergodic theorem ($1 < p < \infty$) for so-called **Markov operators**: If T is a linear operator defined on each of the Banach spaces $L_p(X)$ ($1 \leq p \leq \infty$) by means of the formula $Tf(x) = \int f(y)P(x, dy)$, where $P(x, B)$ is the \dagger transition probability of a \dagger Markov process on

(X, \mathcal{B}) leaving the measure m invariant (i.e., $\int P(x, B) dm = m(B)$ for every $B \in \mathcal{B}$) (\rightarrow 261 Markov Processes), then for every f belonging to $L_p(X)$ ($1 < p < \infty$) the sequence $A_n f$ converges in the norm of $L_p(X)$ to a limit function f^* .

(2) Birkhoff's ergodic theorem has been extended to the following individual ergodic theorem by E. Hopf (1954): If T is a \dagger positive linear operator mapping $L_1(X)$ into $L_1(X)$ and $L_\infty(X)$ into $L_\infty(X)$ with $\|T\|_1 \leq 1$ and $\|T\|_\infty \leq 1$, then for every f in $L_1(X)$ the sequence $A_n f$ converges a.e. to a limit f^* . If T is a Markov operator, then T maps each $L_p(X)$ into itself and satisfies $\|T\|_p \leq 1$ for each p ($1 \leq p \leq \infty$), and therefore Hopf's ergodic theorem applies to such T . Special cases of this theorem were proved earlier by J. Doob and by Kakutani. Later, N. Dunford and J. Schwartz showed that the assumption of the positivity of T can be dispensed with in Hopf's theorem. For a positive linear operator T on $L_1(X)$ satisfying $\|T\|_1 \leq 1$, R. Chacon and D. Ornstein (1960) proved that the **ratio ergodic theorem** holds: For every pair of functions f and g in $L_1(X)$ with $g \geq 0$ a.e., $\lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} T^k f(x) / \sum_{k=0}^{n-1} T^k g(x)$ exists and is finite a.e. on the set $\{x | \sum_{k=0}^{n-1} T^k g(x) > 0\}$. This theorem extends earlier results of Hopf and W. Hurewicz dealing with special classes of operators arising from measurable transformations. Hopf's ergodic theorem can be deduced from the Chacon-Ornstein theorem, while it is known that there are positive operators T on $L_1(X)$ satisfying $\|T\|_1 \leq 1$ for which $\lim_{n \rightarrow \infty} A_n f$ fails to exist on a set of positive measure for some $f \in L_1(X)$. This shows that the assumption $\|T\|_\infty \leq 1$ is crucial in Hopf's theorem.

(3) As was the case in the original proof by Birkhoff of his ergodic theorem, every known proof of an individual ergodic theorem depends crucially on the so-called **maximal ergodic lemma** (or **maximal inequality**). For the case of a positive linear operator T on $L_1(X)$ with $\|T\|_1 \leq 1$, Hopf proved the relevant maximal ergodic lemma: If $E(f)$ is the set $\{x | \sup_{n \geq 1} A_n f(x) > 0\}$ for each f in $L^1(X)$, then $\int_{E(f)} f dm \geq 0$. Hopf's original proof of this lemma was quite intricate, but A. Garcia (1965) succeeded in giving an extremely simple and elegant proof. From the maximal ergodic lemma the following so-called **dominated ergodic theorem** can also be obtained: Let T be a linear operator mapping each $L_p(X)$ into $L_p(X)$ satisfying $\|T\|_p \leq 1$ for each p ($1 \leq p \leq \infty$). If for f in $L_p(X)$ we let $\tilde{f}(x) = \sup_{n \geq 1} |A_n f(x)|$, then if $1 < p < \infty$ we have

$$\int \tilde{f}(x)^p dm \leq \frac{2^p}{p-1} \int |f(x)|^p dm,$$

while if $p = 1$ and $m(X) < \infty$ we have

$$\int \tilde{f}(x) dm \leq 2 \left[m(X) + \int |f(x)| \log^+ |f(x)| dm \right].$$

This theorem was obtained first by N. Wiener (1939) for the special case of T induced by a measure-preserving transformation. M. Akcoglu (1975) proved an isometric "dilation" theorem for positive contractions (i.e., positive linear operators T for which $\|T\| \leq 1$) in the space $L_p(X)$ for $1 < p < \infty$, by means of which he was able to deduce a dominated ergodic theorem for an arbitrary positive contraction in $L_p(X)$ from the corresponding theorem for a positive linear isometry in $L_p(X)$ proved earlier by A. Ionescu-Tulcea. From this theorem of Akcoglu one can obtain the following individual ergodic theorem: If T is a positive contraction in $L_p(X)$ for some p with $1 < p < \infty$, then for every f in $L_p(X)$, the sequence $A_n f$ converges a.e.

(4) Both mean and individual ergodic theorems can be extended without difficulty to a continuous time parameter semigroup $\{T_t | t \geq 0\}$ of bounded linear operators such that $T_t T_s = T_{t+s}$ ($T_0 = I$), under a suitable continuity assumption on T_t with respect to t , by replacing the discrete time average $(\sum_{k=0}^{n-1} T^k)/n$ with $(\int_0^t T_s ds)/t$. Further extensions to n -parameter semigroups were obtained by N. Wiener and by N. Dunford and A. Zygmund. For mean ergodic theorems, even further extensions were possible to \dagger amenable semigroups of bounded linear operators. For 1-parameter semigroups, the behavior of the mean at zero, $(\int_0^t T_s ds)/t$ as $t \downarrow 0$ (**local ergodic theorem**) or $\lambda \int_0^\infty e^{-\lambda s} T_s ds$ as $\lambda \downarrow 0$ or $\lambda \uparrow \infty$ (**Abelian ergodic theorem**), has also been investigated by Wiener, U. Krengel, E. Hille, Yosida, and others. Abelian ergodic theorems are related to properties of the \dagger resolvent of the semigroup $\{T_t\}$ (\rightarrow 378 Semigroups of Operators and Evolution Equations). Further extensions of all these theorems in various directions have been given by many authors; for these extensions and related topics \rightarrow [5–7].

(5) J. F. C. Kingman (1968) proved an interesting and useful extension of (both mean and individual) ergodic theorems, called the **subadditive ergodic theorem**. A real-valued \dagger stochastic process $\{X_{i,k} | 0 \leq i < k, k = 1, 2, 3, \dots\}$ is called a **subadditive process** if it satisfies the following conditions: (i) Whenever $i < j < k$, $X_{i,k} \leq X_{i,j} + X_{j,k}$. (ii) The \dagger joint distribution of $\{X_{i+1,j+1}\}$ is the same as that of $\{X_{i,j}\}$. (iii) The \dagger expectation $g_k = E(X_{0,k})$ exists, and satisfies $g_k \geq -Ak$ for some constant A and for all $k > 1$. Kingman proved that if $\{X_{i,k}\}$ is a subadditive stochastic process, then the limit $\xi = \lim_{n \rightarrow \infty} (1/n)X_{0,n}$ exists a.e. and in the mean,

and $E(\xi) = \lim_{n \rightarrow \infty} (1/n)g_n$. Note that if φ is a measure-preserving transformation on a finite measure space (X, \mathcal{B}, m) and if f is an element of $L_1(X)$, then $X_{i,k} = \sum_{j=i}^{k-1} f \circ \varphi^j$ for $0 \leq i < k$, where $k = 1, 2, 3, \dots$ defines a subadditive process (in fact, $X_{i,k} = X_{i,j} + X_{j,k}$ for $i < j < k$). The **multiplicative ergodic theorem**, proved by V. Oseledec, which plays a significant role in ergodic theory of \dagger nonhyperbolic smooth dynamical systems and which has found applications also in \dagger algebraic groups, can be derived from the subadditive ergodic theorem [8, 9].

C. Recurrence and Invariant Measures

In this section we assume that the measure space (X, \mathcal{B}, m) is nonatomic. A nonsingular measurable transformation φ defined on (X, \mathcal{B}, m) is called **recurrent (infinitely recurrent)** if for every set B and for almost all $x \in B$, there exists an $n \in \mathbb{Z}^+$ (infinitely many $n \in \mathbb{Z}^+$) such that $\varphi^n(x) \in B$. A set W is called **wandering under** φ if $\varphi^{-n}(W) \cap \varphi^{-k}(W) = \emptyset$ for $n \neq k$. The transformation φ is called **conservative** if no sets of positive measure are wandering under φ , and **incompressible** if $B \supset \varphi^{-1}B$ implies $m(B - \varphi^{-1}B) = 0$. The following statements about a nonsingular measurable transformation φ are equivalent: (i) φ is recurrent; (ii) φ is infinitely recurrent; (iii) φ is incompressible; (iv) φ is conservative. An immediate consequence of this is the following **recurrence theorem** of Poincaré (in the form formulated by Carathéodory): A measure-preserving transformation on a finite measure space is infinitely recurrent. In fact, in order for a nonsingular measurable transformation φ to be recurrent (and hence infinitely recurrent) it is sufficient that there exist a finite measure μ invariant under φ and equivalent to (i.e., mutually \dagger absolutely continuous with) the given measure m .

The **invariant measure problem** is one of the basic problems in ergodic theory and is formulated abstractly in the following way: Given a nonsingular measurable transformation φ on a σ -finite measure space (X, \mathcal{B}, m) , find necessary and sufficient conditions for the existence of a finite (or σ -finite) measure invariant under φ and equivalent to m . The given measure m specifies only the class of equivalent measures among which an invariant measure is to be found. Therefore we can assume without loss of generality that m is a finite measure. For the remainder of this section, unless we explicitly state otherwise, we always mean by an invariant measure the one that is equivalent to m .

The Poincaré recurrence theorem states that φ being recurrent is necessary for the existence

of a finite invariant measure. The recurrence of φ is, however, not sufficient, since an ergodic transformation with an infinite but σ -finite invariant measure is recurrent and has no finite invariant measure. A necessary and sufficient condition for the existence of a finite invariant measure was given by a theorem of A. Hajian and Kakutani (1964) which states: φ has a finite invariant measure if and only if φ has no **weakly wandering sets** (a set W is called **weakly wandering under** φ if there exists an infinite subset $\{n_k\}$ of \mathbb{Z}^+ such that $\varphi^{-n_k}W \cap \varphi^{-n_j}W = \emptyset, k \neq j$). Hajian also proved that a bimeasurable transformation φ has a finite invariant measure if and only if φ is **strongly recurrent** in the following sense: For every set E with $m(E) > 0$, there exists a positive integer $k = k(E)$ such that

$$\max_{0 \leq j \leq k} m(\varphi^{n-j}E \cap E) > 0$$

for every $n \in \mathbb{Z}$.

H. Fürstenberg [11] obtained the following striking extension of the Poincaré recurrence theorem: If a bimeasurable transformation φ possesses a finite invariant measure, then for every set E with $m(E) > 0$ and for every integer $k \geq 2$, there exists $n \geq 1$ such that $m(E \cap \varphi^n E \cap \varphi^{2n} E \cap \dots \cap \varphi^{(k-1)n} E) > 0$. From this theorem one can deduce a difficult theorem of E. Szemerédi on \dagger arithmetic progressions which states: Any subset of the integers having positive \dagger upper density contains arithmetic progressions of arbitrary length. In fact, it is not difficult to show that the theorems of Fürstenberg and of Szemerédi are mutually equivalent, and for this reason the theorem of Fürstenberg is sometimes referred to as the **ergodic Szemerédi theorem**.

For a nonsingular bimeasurable transformation φ , a pair of sets A and B are said to be **countably equivalent under** φ if there exist countable decompositions $\{A_k | k \in \mathbb{Z}^+\}$ and $\{B_k | k \in \mathbb{Z}^+\}$ for A and B , respectively, and an infinite subset $\{n_k\}$ of \mathbb{Z} such that $\varphi^{n_k}A_k = B_k$ for each k . A and B are said to be **finitely equivalent under** φ if finite decompositions $\{A_k\}$ and $\{B_k\}$ can be chosen. It was proved by Hopf that (i) φ is recurrent if and only if no set of positive measure is finitely equivalent under φ to one of its proper subsets, and (ii) φ has a finite invariant measure if and only if no set of positive measure is countably equivalent under φ to one of its proper subsets. It can be shown that if φ is ergodic, then φ has no σ -finite invariant measure if and only if every pair of sets of positive measure are countably equivalent under φ .

The first example of a transformation admitting no σ -finite invariant measure was constructed by Ornstein in 1960. Since then, a

number of simpler examples have been obtained by L. Arnold, Brunel, and others. It is now known that there are many different types of transformations having no σ -finite invariant measure (\rightarrow Section F). Furthermore, it was shown by Ionescu-Tulcea that in the group of all nonsingular bimeasurable transformations with a suitable metric, those having a σ -finite invariant measure form a subset of the \dagger first category.

Various extensions of results on the invariant measure problem for nonsingular transformations to the case of \dagger Markov processes having nonsingular transition probabilities were obtained by Y. Ito, J. Neveu, S. Foguel, and others [12].

For investigation of detailed properties of particular transformations arising from classical dynamical systems, problems in number theory, and so on, rather than the existence of invariant measures it is more important to determine a specific form of an invariant measure with desirable properties and to develop methods to decide when such a measure is unique. Various people have considered special classes of transformations; these workers have been able to obtain explicit descriptions of invariant measures with nice properties for the transformations in question and have derived interesting consequences. Most important among these is the so-called **Gibbs measure**, introduced and investigated by Ya. Sinai. He obtained this notion by generalizing the concept of the **equilibrium Gibbs distribution**, which plays a prominent role in \dagger statistical mechanics. It is defined in the following way: Let X be a compact metric space, φ a homeomorphism on X , and μ_0 a probability measure on X invariant under φ . For a function g belonging to $L_\infty(X)$ and for $m, n > 0$ let

$$\Xi_{m,n}(g | \mu_0) = \int_X \exp \left(\sum_{k=-n}^m g(\varphi^k x) \right) d\mu_0(x),$$

and form a sequence of probability measures $\mu_{m,n}(g)$ absolutely continuous with respect to μ_0 , for which the \dagger Radon-Nikodym derivative

$$\frac{d\mu_{m,n}(g | g)}{d\mu_0(x)} = \frac{\exp \sum_{k=-n}^m g(\varphi^k x)}{\Xi_{m,n}(g | \mu_0)}$$

holds. A measure that is a limit point, in the sense of \dagger weak convergence, of the sequence of measures $\mu_{m,n}(g)$ is called a Gibbs measure constructed from μ_0 and g . It is clear that a Gibbs measure is invariant under φ . For mixing topological Markov shifts (\rightarrow Section D), Sinai showed that if one starts with the invariant measure μ_0 of **maximal entropy** (\rightarrow Section H), the existence of which was earlier shown by W. Parry, one can determine a class of functions g for which the Gibbs measure

is unique, i.e., $\mu(g) = \lim_{m,n \rightarrow \infty} \mu_{m,n}(g)$ exists. Sinai further investigated the properties of the unique measure $\mu(g)$ in detail. R. Bowen and D. Ruelle, defining the Gibbs measure somewhat differently, investigated the existence and uniqueness of such measures, thereby recapturing the results of Sinai in the case of mixing topological Markov shifts. With the aid of **Markov partitions** (\rightarrow Section G), these results on Gibbs measures were carried over to the case of \dagger Anosov and “axiom A” diffeomorphisms, and they provided essential tools in the investigation of the ergodic behavior of these transformations. For details \rightarrow [13, 14].

The explicit form of the density of the invariant measure with respect to the Lebesgue measure for the transformation associated with \dagger continued fraction expansion was already known to Gauss, and one can draw from it numerous conclusions about metric properties of continued fractions. The way in which Gauss was able to determine this invariant measure, however, was never explained. Recently, Sh. Ito, H. Nakada, and S. Tanaka (*Keio Eng. Rep.*, 30 (1977)) developed an interesting method to describe the mechanism used to arrive at this density function for the invariant measure for the continued fraction transformation. They employed similar methods in subsequent work to determine explicit density functions for invariant measures for other related number-theoretic transformations and for certain classes of continuous mappings over an interval; by means of the explicit forms of invariant measures they were able to describe the metric properties of these transformations in detail.

D. Examples and Construction of Measure-Preserving Transformations

Examples of measure-preserving transformations appear in many different contexts. We describe some of the important ones.

(1) Let G be a \dagger locally compact Abelian group satisfying the second axiom of \dagger countability (\rightarrow 423 Topological Groups), \mathcal{B} a σ -algebra of Borel subsets of G , and m its \dagger Haar measure (normalized if G is compact). Then (G, \mathcal{B}, m) is a Lebesgue measure space. For a fixed element $g_0 \in G$, define the transformation $\varphi_{g_0}: G \rightarrow G$ by $\varphi_{g_0}(g) = g + g_0$. Then φ_{g_0} is a bijective measure-preserving transformation on (G, \mathcal{B}, m) and is called the **rotation** on G by the element g_0 . If G is compact, then the rotation φ_{g_0} is ergodic if and only if the cyclic subgroup generated by the element g_0 is dense in G . If this happens, the element g_0 is called the **topological generator** of G . A group

is called **monothetic** if it has a topological generator.

(2) If φ is a group \dagger endomorphism of a compact Abelian group G , then φ preserves the Haar measure m . If φ is a group \dagger automorphism, then it is a bijective measure-preserving transformation on (G, \mathcal{B}, m) . A continuous group automorphism φ induces a group automorphism φ^* of the character group G^* .

The measure-preserving transformation φ is ergodic if and only if every character except the identity has an infinite orbit under the induced automorphism φ^* . When the group is the n -dimensional torus \mathbf{T}^n , a continuous group automorphism φ is uniquely represented by an $n \times n$ matrix with integer entries and with determinant ± 1 . In this case, φ is ergodic if and only if no roots of unity appear among the eigenvalues of the representing matrix.

(3) Let (Y, \mathcal{A}) be a measurable space, let $(Y_n, \mathcal{A}_n) = (Y, \mathcal{A})$ for each $n \in \mathbb{Z}$, and define (Y^*, \mathcal{A}^*) to be the \dagger product measurable space $(\prod_{n \in \mathbb{Z}} Y_n, \prod_{n \in \mathbb{Z}} \mathcal{A}_n)$. The transformation φ defined on (Y^*, \mathcal{A}^*) by

$$y^* = (\dots, y_{-1}, y_0, y_1, \dots) \\ \rightarrow \varphi(y^*) = (\dots, y'_{-1}, y'_0, y'_1, \dots)$$

with $y'_n = y_{n+1}$ for each n is called the **shift transformation**. Let μ be a probability measure on (Y, \mathcal{A}) such that (Y, \mathcal{A}, μ) is a Lebesgue measure space, let $\mu_n = \mu$ for each n , and define μ^* to be the \dagger product measure $\prod_{n \in \mathbb{Z}} \mu_n$ on (Y^*, \mathcal{A}^*) . Then $(Y^*, \mathcal{A}^*, \mu^*)$ is a Lebesgue measure space, and the shift transformation φ is a bijective measure-preserving transformation. Considered with the product measure, φ is called a **generalized Bernoulli shift**. When the set Y is at most countable and the measure μ on Y is given by a sequence $\{p_j\}$ of positive numbers with $\sum p_j = 1$, φ is called a **Bernoulli shift**. Suppose that $P(y, A)$ is a Markov transition function on (Y, \mathcal{A}) and π is a probability measure invariant under $P(y, A)$. Then we can define the **Markov measure** π^* on the product space (Y^*, \mathcal{A}^*) by setting

$$\pi^*(E^*) = \int_{A_{i+s}} \dots \int_{A_{i+1}} \int_{A_i} \pi(dy_0) P(y_0, dy_1) \\ \times P(y_1, dy_2) \dots P(y_{s-1}, dy_s),$$

for a cylinder set

$$E^* = \left(\prod_{j < i} Y_j \right) \times \left(\prod_{i \leq k \leq i+s} A_k \right) \times \left(\prod_{n > i+s} Y_n \right),$$

and extending it to all of \mathcal{A}^* . The shift transformation φ preserves the Markov measure π^* . Considered as a measure-preserving transformation on $(Y^*, \mathcal{A}^*, \pi^*)$, φ is called a **Markov shift**.

A generalized Bernoulli shift is always ergodic. A Markov shift is ergodic if and only if the corresponding Markov process is irreducible, which is the case if and only if the following property is satisfied: For every pair of sets A and B with $\pi(A)\pi(B) > 0$, there exists an $n \in \mathbb{Z}^+$ such that $\int_A P^n(y, B) d\pi > 0$.

There are other measures besides the product measure and the Markov measure that can be defined on the product space (Y^*, \mathcal{A}^*) and are invariant under the shift transformation φ . For example, if Y is a Borel subset of \mathbb{R} and \mathcal{A} is the σ -algebra of Borel subsets of Y , then any \dagger stationary stochastic process taking values in Y induces such a measure on (Y^*, \mathcal{A}^*) . When considered with a measure of this type, the shift transformation φ is called the **shift associated with the stationary process**. Properties of the shift associated with a \dagger stationary Gaussian process have been investigated by G. Maruyama, I. Girsanov, H. Totoki, and others. In particular, it is known that the shift is ergodic if and only if the \dagger spectral measure for the \dagger covariance function of the associated Gaussian process is continuous [15].

(4) Other important examples of measure-preserving transformations arise from classical dynamical systems, which will be described in Section G.

(5) There are several ways of constructing new measure-preserving transformations from given ones. We describe important cases.

(i) Let φ be a nonsingular, measurable, recurrent transformation (not necessarily measure-preserving) on a σ -finite measure space (X, \mathcal{B}, m) , and let A be a set of positive measure. For $x \in A$, let $n(x) = \min\{n \in \mathbb{Z}^+ \mid \varphi^n(x) \in A\}$. The transformation $\varphi_A: A \rightarrow A$ defined by $\varphi_A(x) = \varphi^{n(x)}(x)$ is a nonsingular measurable transformation on the measure space $(A, \mathcal{B} \cap A, m_A)$, where $m_A(B) = m(A \cap B)/m(A)$, and it is measure-preserving if φ is. We call φ_A the **transformation induced by φ on A** . It is ergodic if φ is ergodic.

(ii) Let φ be a nonsingular measurable transformation on a σ -finite measure space (X, \mathcal{B}, m) , and suppose that $\{A_n\}$ is a countable (possibly finite) partition of X . Define a function $f: X \rightarrow \mathbb{Z}^+$ by setting $f(x) = n$ for $x \in A_n$, and let $\tilde{X} = \{(x, j) \mid x \in X, 1 \leq j \leq f(x)\}$ be a subspace of the product measure space $(X \times \mathbb{Z}^+, \mathcal{B} \times \mathcal{C}, m \times \mu)$, where \mathcal{C} is the σ -algebra of all subsets of \mathbb{Z}^+ , and μ is the measure on $(\mathbb{Z}^+, \mathcal{C})$ defined by $\mu(\{n\}) = 1$ for each $n \in \mathbb{Z}^+$. The transformation $\tilde{\varphi}$ defined on \tilde{X} by $\tilde{\varphi}(x, j) = (x, j+1)$ if $1 \leq j < f(x)$ and $= (\varphi(x), 1)$ if $j = f(x)$ is a nonsingular measurable transformation on the measure space $(\tilde{X}, (\mathcal{B} \times \mathcal{C}) \cap \tilde{X}, (m \times \mu)_{\tilde{X}})$, and it is measure-preserving if φ is. We

call $\tilde{\varphi}$ the **transformation built from φ with the ceiling function f** . If φ is ergodic and $m(A_n) \rightarrow 0$ as $n \rightarrow \infty$, then $\tilde{\varphi}$ is also ergodic.

(iii) Suppose that ψ is a measure-preserving transformation on (X, \mathcal{B}, m) and φ_x is a measure-preserving transformation on (Y, \mathcal{A}, μ) for each $x \in X$. Assume that the mapping $(x, y) \rightarrow \varphi_x(y)$ is measurable with respect to the σ -algebras $\mathcal{B} \times \mathcal{A}$ and \mathcal{A} . The transformation θ defined on the product space $(X \times Y, \mathcal{B} \times \mathcal{A}, m \times \mu)$ by $\theta(x, y) = (\psi(x), \varphi_x(y))$ is measure-preserving and is called the **skew product** of ψ and $\{\varphi_x\}$. If $\varphi_x = \varphi$ for all $x \in X$, then we get a **direct product** transformation $\theta(x, y) = (\psi(x), \varphi(y))$.

(iv) A measure-preserving transformation φ on (Y, \mathcal{A}, μ) is said to be a **factor transformation** (or a **homomorphic image**) of a measure-preserving transformation ψ on (X, \mathcal{B}, m) if there exists a measurable transformation η from X onto Y such that $m \circ \eta^{-1} = \mu$ and $\varphi \eta = \eta \psi$. If \mathcal{B}' is a σ -subalgebra of \mathcal{B} and ψ leaves \mathcal{B}' invariant (i.e., $\psi^{-1}\mathcal{B}' \subset \mathcal{B}'$), then ψ induces a factor transformation φ on the measure space (X, \mathcal{B}', m) . Conversely, if a measure-preserving transformation φ on (Y, \mathcal{A}, μ) is a factor transformation of ψ on (X, \mathcal{B}, m) via a mapping η , then $\mathcal{B}' = \eta^{-1}\mathcal{A}$ is a σ -subalgebra of \mathcal{B} invariant under ψ .

(6) A one-parameter family $\{\varphi_t \mid t \in \mathbb{R}\}$ of bijective measure-preserving transformations on a measure space (X, \mathcal{B}, m) is called a **flow**. A flow is called **continuous** if the mapping $t \rightarrow T_t$ is \dagger weakly continuous where $\{T_t\}$ is the one-parameter family of \dagger unitary operators on $L^2(X)$ induced by the flow $\{\varphi_t\}$. A flow is called **measurable** if the mapping $(t, x) \rightarrow \varphi_t(x)$ is a measurable transformation of $\mathbb{R} \times X$ into X . A measurable flow is continuous. A. Vershik and Maruyama proved that for any continuous flow there exists a measurable flow, unique in a specified sense, which is spatially isomorphic (in the sense specified in Section E) to the given flow.

Important examples of flows are given by classical dynamical systems (\rightarrow Section G), and by continuous-time stationary stochastic processes.

An important tool in the study of flows is provided by the theorem of W. Ambrose and Kakutani: Every measurable ergodic flow without a fixed point is spatially isomorphic to an **S-flow**. A measurable flow $\{\varphi_t\}$ is called an **S-flow** (**special flow** or **flow built under a function**) if there exist a measure-preserving transformation φ of a measure space (X, \mathcal{B}, m) and an \mathbb{R}^+ -valued function f on (X, \mathcal{B}, m) such that each φ_t is a measure-preserving transformation on the subspace $\tilde{X} = \{(x, u) \mid x \in X, 0 \leq u \leq f(x)\}$ of the product measure space

$(X \times \mathbf{R}^+, \mathcal{B} \times \mathcal{M}, m \times \lambda)$ given by

$$\begin{aligned} \varphi_t(x, u) &= (x, u + t) \quad \text{if} \quad -u \leq t < -u + f(x), \\ &= \left(\varphi^n(x), u + t - \sum_{k=0}^{n-1} f(\varphi^k(x)) \right) \quad \text{if} \\ &\quad -u + \sum_{k=0}^{n-1} f(\varphi^k(x)) \leq t < -u + \sum_{k=0}^n f(\varphi^k(x)), \\ &= \left(\varphi^{-n}(x), u + t + \sum_{k=-n}^{-1} f(\varphi^k(x)) \right) \quad \text{if} \\ &\quad -u - \sum_{k=-n}^{-1} f(\varphi^k(x)) \leq t < -u - \sum_{k=-n+1}^{-1} f(\varphi^k(x)), \end{aligned}$$

$n \geq 1$. Here \mathcal{M} is the σ -algebra of Borel subsets of \mathbf{R}^+ , and λ is the usual Lebesgue measure.

E. Isomorphism Problems

In this section we assume that the Lebesgue measure spaces considered are probability spaces. For simplicity, following the common usage among Russian mathematicians, we call a measure-preserving transformation on (X, \mathcal{B}, m) an **endomorphism** and a bijective measure-preserving transformation an **automorphism**.

An automorphism φ_1 (a flow $\{\varphi_1^{(t)}\}$) on $(X_1, \mathcal{B}_1, m_1)$ is said to be **spatially isomorphic** (or **metrically isomorphic**) to an automorphism φ_2 (a flow $\{\varphi_2^{(t)}\}$) on $(X_2, \mathcal{B}_2, m_2)$ if there exist sets N_1 and N_2 with $m_1(N_1) = m_2(N_2) = 0$ and a bijective measurable transformation θ from $X_1 - N_1$ to $X_2 - N_2$ such that $m_2 \circ \theta = m_1$ and $\theta \varphi_1 = \varphi_2 \theta$ ($\theta \varphi_1^{(t)} = \varphi_2^{(t)} \theta$ for each t).

Classification of automorphisms and flows into isomorphism classes constitutes the central problem of modern ergodic theory. Properties of automorphisms and flows that are preserved under spatial isomorphisms are called **isomorphism invariants** (or **metric invariants**). There are several isomorphism invariants that are essential to the study of the isomorphism problem. We describe these below for the case of automorphisms. There are corresponding invariants for flows as well.

(1) Spectral Invariants. Two automorphisms φ_1 and φ_2 are said to be **spectrally isomorphic** if the unitary operators T_1 and T_2 induced by φ_1 and φ_2 on the Hilbert spaces $L_2(X_1)$ and $L_2(X_2)$, respectively, are unitarily equivalent (i.e., there exists an isometric isomorphism V of $L_2(X_1)$ onto $L_2(X_2)$ such that $VT_1 = T_2V$). Properties preserved under spectral isomorphisms are called **spectral invariants** (or **spectral properties**). If φ_1 and φ_2 are spatially isomorphic, it is clear that they are spectrally iso-

morphic; but the converse is not true in general.

(i) The property of φ being ergodic is a spectral property since φ is ergodic if and only if the number 1 is a simple eigenvalue of the induced unitary operator T . If φ is ergodic, then the set of all eigenvalues of the induced operator T forms a subgroup of the circle group, each eigenvalue is simple, and each eigenfunction has constant absolute value. If the spectrum of the induced operator T consists entirely of eigenvalues, φ is said to have **discrete spectrum** (or **pure point spectrum**). A theorem due to von Neumann and P. Halmos—the first theorem on the question of isomorphism—states that two ergodic automorphisms φ_1 and φ_2 with discrete spectra are spatially isomorphic if and only if they are spectrally isomorphic, which is the case if and only if the induced operators T_1 and T_2 have the same set of eigenvalues. Furthermore, every ergodic automorphism with discrete spectrum is spatially isomorphic to an ergodic rotation on a compact Abelian group [2].

Analogous results were obtained by L. Abramov for a bigger class of automorphisms, namely, for ergodic automorphisms having so-called **quasidiscrete spectra**.

(ii) An automorphism φ is ergodic if and only if for every pair of sets A, B ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{k=0}^{n-1} m(\varphi^k(A) \cap B) \right) = m(A)m(B).$$

Strengthening this condition, we can define φ to be **weakly mixing** if for every pair of sets A, B ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{k=0}^{n-1} |m(\varphi^k(A) \cap B) - m(A)m(B)| \right) = 0;$$

strongly mixing if

$$\lim_{k \rightarrow \infty} m(\varphi^k(A) \cap B) = m(A)m(B);$$

and **k -fold mixing** if for arbitrary choice of sets $A_j, j = 0, 1, \dots, k$,

$$\begin{aligned} \lim m(A_0 \cap \varphi^{n_1} A_1 \cap \varphi^{n_2} A_2 \dots \cap \varphi^{n_k} A_k) \\ = m(A_0)m(A_1) \dots m(A_k), \end{aligned}$$

where the limit is taken as $n_1, n_2, \dots, n_k \rightarrow \infty$ in such a way that $n_1 < n_2 < \dots < n_k$ and $\min_{1 \leq j \leq k} (n_j - n_{j-1}) \rightarrow \infty$. The property of an automorphism φ being weakly mixing or strongly mixing is a spectral property. For instance, φ is weakly mixing if and only if the number 1 is a simple eigenvalue and is the only eigenvalue of the induced operator T . It is also known that φ is weakly mixing if and only if the direct product automorphism $\varphi \times \varphi$ is ergodic. The set of all weakly mixing automorphisms forms a dense G_δ -set in the group

of all automorphisms on (X, \mathcal{B}, m) considered with the so-called weak topology (Halmos's theorem). On the other hand, it was shown by V. Rokhlin that the set of all strongly mixing automorphisms is a set of first category with respect to the weak topology. However, there are only a few known examples of automorphisms that are weakly mixing but not strongly mixing.

(iii) An automorphism φ is said to have **countable Lebesgue spectrum** if the maximal spectral type of the induced unitary operator T restricted to the $^\perp$ orthocomplement of the subspace of constant functions in $L_2(X)$ is equivalent to the Lebesgue measure and its $^\perp$ multiplicity is countably infinite.

(iv) An automorphism φ is called a **K-automorphism** (or **Kolmogorov automorphism**) if there exists a σ -subalgebra \mathcal{B}_0 of \mathcal{B} such that (a) $\varphi\mathcal{B}_0 \supset \mathcal{B}_0$ and $\varphi\mathcal{B}_0 \neq \mathcal{B}_0$, (b) $\bigvee_{n \in \mathbb{Z}} \varphi^n \mathcal{B}_0 = \mathcal{B}$, and (c) $\bigwedge_{n \in \mathbb{Z}} \varphi^n \mathcal{B}_0 = \mathcal{N}$, where \mathcal{N} is the σ -subalgebra of \mathcal{B} consisting of null sets and their complements. The notion of a **K-flow** (or **Kolmogorov flow**) is defined similarly. K-automorphisms are k -fold mixing for all orders k and have countable Lebesgue spectra.

Generalized Bernoulli shifts are all K-automorphisms. An ergodic Markov shift is a K-automorphism if and only if it is strongly mixing, which is the case if and only if the corresponding Markov process is $^\perp$ irreducible and $^\perp$ aperiodic (\rightarrow 260 Markov Chains B). A continuous group automorphism of a compact Abelian group is a K-automorphism if and only if it is ergodic. In particular, a continuous group automorphism of the n -dimensional torus \mathbb{T}^n is a K-automorphism if and only if no roots of unity appear among the eigenvalues of the representing matrix. Automorphisms and flows arising from classical dynamical systems also provide examples of K-automorphisms and K-flows. In particular, a $^\perp$ geodesic flow on a surface of negative curvature is a K-flow, and each automorphism (except the identity) of this flow is a K-automorphism. For the shift transformation φ associated with a stationary Gaussian process, it was shown by Maruyama [15] that φ is (a) weakly mixing if and only if it is ergodic, (b) strongly mixing if and only if the covariance function of the associated Gaussian process tends to 0 as $n \rightarrow \infty$, and (c) a K-automorphism if and only if the $^\perp$ spectral measure of the covariance function is absolutely continuous with respect to the Lebesgue measure.

(v) Examples of automorphisms having various types of spectra have been constructed by a number of authors by using stationary Gaussian processes and the theory of approximation developed by A. Katok and A. Stepin [16].

(vi) An ergodic automorphism with quasi-discrete spectrum has a mixed spectrum, that is, the spectrum of the induced unitary operator T has a continuous component and eigenvalues in addition to 1. Anzai (1951) constructed a special class of skew product automorphisms having mixed spectra and showed that in this class there are automorphisms that are spectrally isomorphic but not spatially isomorphic. However, the question of whether two spatially nonisomorphic automorphisms exist among automorphisms having the same purely continuous spectrum remained unanswered for a long time, until in 1958 Kolmogorov (*Dokl. Akad. Nauk SSSR*, (5) 119 (1958)) settled it affirmatively by using a new isomorphism invariant called **entropy**.

2. Generators and Entropy. (i) By a **partition** $\xi = \{A_\lambda\}$ of the space X we mean a collection of sets A_λ such that $A_\lambda \cap A_{\lambda'} = \emptyset$ whenever $\lambda \neq \lambda'$ and $\bigcup A_\lambda = X$. We denote by ε the partition of X into individual points, and by ν the trivial partition $\{X\}$. A partition into a finite (countable) number of sets is called a finite (countable) partition. A partition ξ is said to be finer than another partition ζ (or ζ is coarser than ξ) if for every $A \in \xi$ there is a set $B \in \zeta$ such that $A \subset B$. For a collection $\{\xi_\alpha\}$ of partitions of X , we denote by $\bigvee_\alpha \xi_\alpha$ the coarsest partition that is finer than each ξ_α , and by $\bigwedge_\alpha \xi_\alpha$ the finest partition that is coarser than each ξ_α . If $\xi_k = \{A_{k,n}\}$ is a sequence of countable (or finite) partitions, then $\bigvee_k \xi_k$ is precisely the partition of X into nonempty intersections of the form $\bigcap_k A_{k,n_k}$ with $A_{k,n_k} \in \xi_k$ for each k . With a partition ξ of X we associate a σ -subalgebra $\mathcal{B}(\xi)$ of \mathcal{B} which is the σ -algebra of all \mathcal{B} -measurable sets that are a union of elements in ξ . Two partitions ξ and ζ are said to coincide a.e. if $\mathcal{B}(\xi) = \mathcal{B}(\zeta)$ a.e. (i.e., for every $A \in \mathcal{B}(\xi)$ there exists a set B in $\mathcal{B}(\zeta)$ such that $m(A \cup B - A \cap B) = 0$ and conversely).

(ii) Suppose that φ is an endomorphism of (X, \mathcal{B}, m) and ξ a partition of X . By $\varphi^{-1}\xi$ we mean the partition $\{\varphi^{-1}(A) \mid A \in \xi\}$. If φ is an automorphism, we also define $\varphi\xi = \{\varphi(A) \mid A \in \xi\}$. A partition ξ is called a **generator** for an endomorphism φ if $\bigvee_{n=0}^\infty \varphi^{-n}\xi = \varepsilon$ a.s. If $\bigvee_{n=-\infty}^\infty \varphi^n\xi = \varepsilon$ a.s., ξ is called a **two-sided generator** for an automorphism φ .

An endomorphism φ is said to be **periodic at a point** $x \in X$ if there exists a positive integer n such that $\varphi^n(x) = x$ and **aperiodic** if the set of points of periodicity has measure zero. If the measure space (X, \mathcal{B}, m) is nonatomic, then every ergodic endomorphism on it is aperiodic. A theorem of Rokhlin states that every aperiodic automorphism φ has a countable two-sided generator. This implies that every such φ is spatially isomorphic to the shift

transformation on the infinite product space (Y^*, \mathcal{A}^*) considered with some invariant measure μ^* , where each coordinate space $Y_j = Y$ has at most a countable number of points. Krieger improved this result by showing that if an ergodic automorphism φ has finite entropy, then φ has a finite two-sided generator [17].

(iii) For a finite or countable partition $\xi = \{A_n\}$, define the **entropy $H(\xi)$ of the partition** to be $-\sum_n m(A_n) \log(m(A_n))$ (the logarithms here and below are natural logarithms). We denote by \mathcal{L} the set of all partitions ξ with $H(\xi) < \infty$. If $\xi \in \mathcal{L}$, then for any endomorphism φ the limit

$$h(\varphi, \xi) = \lim_{n \rightarrow \infty} \frac{1}{n} \left(H \left(\bigvee_{k=0}^{n-1} \varphi^{-k} \xi \right) \right)$$

exists and is finite. The **entropy $h(\varphi)$ of the endomorphism φ** is defined to be $\sup \{h(\varphi, \xi) \mid \xi \in \mathcal{L}\}$ and is an isomorphism invariant. Properties of entropy have been investigated extensively since the notion was introduced by Kolmogorov. We cite a few results.

(a) If a partition $\xi \in \mathcal{L}$ is a generator for an endomorphism φ or a two-sided generator for an automorphism φ , then $h(\varphi) = h(\varphi, \xi)$ (Sinaï's lemma). (b) For every integer n , $h(\varphi^n) = |n| h(\varphi)$, and for a measurable flow $\{\varphi_t\}$, $h(\varphi_t) = |t| h(\varphi_1)$ for every real number t . (c) If an automorphism φ is periodic, then $h(\varphi) = 0$. (d) If φ_1 is a factor transformation of φ_2 , then $h(\varphi_1) \leq h(\varphi_2)$. (e) $h(\varphi_1 \times \varphi_2) = h(\varphi_1) + h(\varphi_2)$. (There is a more complicated formula (due to Rokhlin) for the entropy of a skew product automorphism.) (f) If φ is a recurrent automorphism and φ_A is the automorphism induced by φ on a subset A with $m(A) > 0$, then $h(\varphi_A) = h(\varphi)/m(A)$ (**Abramov's formula**). (g) If φ is a Bernoulli shift with probability distribution $\{p_n\}$, then $h(\varphi) = -\sum_n p_n \log p_n$. (h) If φ is a Markov shift based on the Markov transition probability P_{ij} (defined on a countable or finite state space) and an invariant measure π_i , then

$$h(\varphi) = -\sum_i \sum_j \pi_i P_{ij} \log P_{ij}.$$

(i) If φ is ergodic and has a pure point spectrum, or more generally, has a quasisdiscrete spectrum, then $h(\varphi) = 0$. (j) For an ergodic group automorphism φ on an n -dimensional torus, $h(\varphi) = \sum \log |\lambda|$, where the sum is taken over all eigenvalues λ of modulus > 1 of the representing matrix. (k) If an automorphism φ has positive entropy, then in $L_2(X)$ there exists a subspace invariant under the induced unitary operator T such that the spectrum of T restricted to this subspace is countable Lebesgue (Rokhlin's theorem). It follows from (k) that automorphisms with singular spectra or spectra of finite multiplicity must have zero entropy. In proving assertion (g), Kolmogorov

established for the first time the fact that there are uncountably many spatially nonisomorphic Bernoulli shifts.

(iv) An automorphism φ is said to have **completely positive entropy** if $h(\varphi, \xi) > 0$ for every partition $\xi \neq v$. It was shown by Rokhlin and Sinaï that an automorphism φ has completely positive entropy if and only if φ is a K -automorphism. M. Pinsker proved that for every automorphism φ there exists a partition, called the **Pinsker partition**, that is invariant under φ and such that the factor transformation of φ with respect to this partition has zero entropy and is the largest among the factor transformations of φ with zero entropy.

(v) Rokhlin showed that $h(\varphi) = 0$ for an endomorphism φ if and only if all of its factor transformations are automorphisms. An endomorphism φ is called **exact** if $\bigwedge_{n=0}^{\infty} \varphi^{-n} \varepsilon = v$ a.e. Rokhlin introduced a way to associate with each endomorphism a certain automorphism, called the **natural extension**, which reflects the properties of the endomorphism. For example, an endomorphism and its natural extension are simultaneously ergodic or nonergodic, are mixing of the same order, and have equal entropy. The natural extension of an exact endomorphism is a K -automorphism.

(vi) Automorphisms φ_1 and φ_2 are said to be **weakly isomorphic** if each of them is a factor transformation of the other. Sinaï proved (1964) that for each ergodic automorphism with positive entropy, there exists a factor automorphism having the same entropy and isomorphic to a Bernoulli shift, and hence it follows in particular that Bernoulli shifts with the same entropy are weakly isomorphic. Ornstein (*Adv. in Math.*, 4 (1970)) went further and succeeded in proving the following remarkable result: Two Bernoulli shifts with equal entropy are spatially isomorphic. Partial results in this direction were obtained earlier by L. Meshalkin, J. Blum, and D. Hanson. In the proof of Ornstein's theorem, essential use was made of the following theorem of C. Shannon and B. McMillan, which plays a fundamental role in information theory (\rightarrow 213 Information Theory): Suppose that φ is an ergodic endomorphism on (X, \mathcal{B}, m) and ξ is a partition of X . For a point $x \in X$, let $A_n(x)$ denote the element in the partition $\bigvee_{k=0}^{n-1} \varphi^{-k} \xi$ that contains x . Then for almost all x ,

$$\lim_{n \rightarrow \infty} \left(-\frac{1}{n} \log m(A_n(x)) \right)$$

exists and equals $h(\varphi, \xi)$.

(vii) Techniques and ideas developed by Ornstein in his proof of the isomorphism theorem have been refined and extended further by himself, B. Weiss, N. Friedman, M. Smorodinsky, and others, and numerous re-

sults were subsequently obtained on the isomorphism problem. We describe below some of the main results and basic concepts; for more detailed accounts \rightarrow [21–24].

A sequence $\{\xi_n\}$ of partitions is said to be **independent** if for every choice of sets A_1, \dots, A_s with $A_k \in \xi_{n_k}$, $m(\bigcap_{k=1}^s A_k) = \prod_{k=1}^s m(A_k)$ whenever n_1, n_2, \dots, n_s are all distinct. For a fixed $\varepsilon > 0$, two partitions ξ and ζ are said to be ε -**independent** if

$$\sum_{A \in \xi, B \in \zeta} |m(A \cap B) - m(A)m(B)| < \varepsilon.$$

A pair (φ, ξ) , where φ is an automorphism of (X, \mathcal{B}, m) and ξ is a finite (or countable) partition of X , is called a **process** (on X).

$\mathcal{B}(\bigvee_{n=-\infty}^{\infty} \varphi^n \xi)$ is a σ -subalgebra of \mathcal{B} invariant under φ , and φ restricted to this σ -subalgebra is a factor automorphism of φ and is isomorphic to a shift transformation on an infinite product space, where each coordinate space has the same number of points as the number of atoms in ξ . A process (φ, ξ) is called a **Bernoulli process** (or an **independent process**) if the sequence of partitions $\{\varphi^n \xi | n \in \mathbb{Z}\}$ is independent, and a **weak Bernoulli (W.B.) process** if for every $\varepsilon > 0$ there exists a $k > 0$ such that the partitions $\bigvee_{i=k}^{k+n} \varphi^i \xi$ and $\bigvee_{i=-n}^0 \varphi^i \xi$ are ε -independent for all $n \geq 0$. Let J be a finite set and α, β be two functions from the set $\{1, 2, \dots, N\}$ to J . By $d_N(\alpha, \beta)$ we denote $(1/N) \# \{n | \alpha(n) \neq \beta(n)\}$, where $\#$ denotes the cardinality. d_N defines a metric on the space $J^{\{1, 2, \dots, N\}}$ called the **Hamming distance**. Now,

for a process (φ, ξ) on X with $\xi = \{A_j | j \in J\}$, we define for $x, y \in X$, $d_N^{\xi}(x, y)$ to be equal to $d_N(\xi_{\varphi}^N(x), \xi_{\varphi}^N(y))$, where for $x \in X$, $\xi_{\varphi}^N(x)$ is the point $(j(x), j(\varphi(x)), \dots, j(\varphi^{N-1}(x))) \in J^{\{1, 2, \dots, N\}}$ and $j(x)$ is the index $j \in J$ for which $x \in A_j$. A process (φ, ξ) is called **very weak Bernoulli (V.W.B.)** if for any $\varepsilon > 0$ there exists an N_0 such that whenever $N > N_0$ there is a set G with $m(G) > 1 - \varepsilon$ belonging to $\mathcal{B}(\bigvee_{n=-\infty}^0 \varphi^n \xi)$ satisfying the following condition: for any $A \in \mathcal{B}(\bigvee_{n=-\infty}^0 \varphi^n \xi)$ with $A \subset G$ and $m(A) > 0$, there exists a probability measure ν on $X \times X$ satisfying (i) $\nu(E \times X) = m(E)$ and $\nu(X \times F) = m(F)$ for all $E, F \in \mathcal{B}$ and (ii) $\int_{X \times X} d_N^{\xi}(x, y) d\nu < \varepsilon$. It is not difficult to show that W.B. processes are V.W.B. For processes (φ, ξ) on (X, \mathcal{B}, m) and $(\bar{\varphi}, \bar{\xi})$ on $(\bar{X}, \bar{\mathcal{B}}, \bar{m})$, where both ξ and $\bar{\xi}$ are indexed by the same finite set J , we define $\bar{d}_N((\varphi, \xi), (\bar{\varphi}, \bar{\xi}))$ to be $\inf \{ \int_{X \times \bar{X}} d_N^{\xi}(x, \bar{x}) d\rho(x, \bar{x}) \}$, where \inf is taken over all probability measures ρ on $X \times \bar{X}$ satisfying $\rho(E \times \bar{X}) = m(E)$ for all $E \in \mathcal{B}$ and $\rho(X \times \bar{E}) = \bar{m}(\bar{E})$ for all $\bar{E} \in \bar{\mathcal{B}}$. \bar{d}_N decreases with N , and we define $\bar{d}((\varphi, \xi), (\bar{\varphi}, \bar{\xi}))$ to be $\lim_{N \rightarrow \infty} \bar{d}_N((\varphi, \xi), (\bar{\varphi}, \bar{\xi}))$. Finally, a process (φ, ξ) is said to be **finitely determined (F.D.)** if for any $\varepsilon > 0$ there exists a $\delta > 0$ and N such that if $(\bar{\varphi}, \bar{\xi})$ is any

process satisfying (i) ξ and $\bar{\xi}$ are indexed by the same set J , (ii) $h(\varphi, \xi) - h(\bar{\varphi}, \bar{\xi}) < \delta$, and (iii) $\bar{d}_N((\varphi, \xi), (\bar{\varphi}, \bar{\xi})) < \delta$, then $\bar{d}((\varphi, \xi), (\bar{\varphi}, \bar{\xi})) < \varepsilon$. Let us call a partition ξ an **F.D. generator** for an automorphism φ if ξ is a two-sided generator for φ and the process (φ, ξ) is finitely determined. A general isomorphism theorem proved by Ornstein states: If automorphisms φ and $\bar{\varphi}$ have the same entropy and both have F.D. generators, then φ and $\bar{\varphi}$ are isomorphic. Ornstein and Weiss proved further that the process (φ, ξ) is F.D. if and only if it is V.W.B. From these basic results a number of results can be deduced. (a) Nontrivial factors of Bernoulli shifts are Bernoulli. (b) Strongly mixing Markov shifts have two-sided weak Bernoulli generators and hence are isomorphic to Bernoulli shifts. (c) Every Bernoulli shift can be embedded in a flow, which implies among other things that Bernoulli shifts have roots of all orders. (d) It was shown by Y. Katznelson that every ergodic group automorphism of an n -dimensional torus has a two-sided generator which can be shown to be V.W.B., and hence is also isomorphic to a Bernoulli shift. R. Adler and B. Weiss earlier proved by entirely different methods that on a 2-dimensional torus, two ergodic group automorphisms having the same entropy are spatially isomorphic. They have done this by constructing a two-sided generator ξ for such automorphisms which is also a Markov partition (\rightarrow Section G).

(viii) Among further positive results there are the following: (a) Except for a trivial change in the time scale, any two Bernoulli flows are isomorphic (Ornstein). (b) Every ergodic group automorphism of a compact group is isomorphic to a Bernoulli shift (Thomas and Miles; Lind). (c) A number of automorphisms and flows arising from classical dynamical systems are shown to be Bernoulli (\rightarrow Section G). (d) Examples of exact endomorphisms arise in connection with problems in number theory; for example, \dagger continued fraction expansion and β -expansion. The natural extensions of the exact endomorphisms associated with continued fraction expansion and β -expansion are now known to be spatially isomorphic to Bernoulli shifts.

(ix) Since many of the automorphisms that have been shown to be K -automorphisms are now known to be isomorphic to Bernoulli shifts, it is natural to expect that every K -automorphism is in fact Bernoulli. However, Ornstein (1973) constructed an example of a K -automorphism that is not a Bernoulli shift. A lot of work has since been done to investigate how bad K -automorphisms can be. It turns out that K -automorphisms share almost none of the finer properties of Bernoulli shifts. For example: (a) There are uncountably many

nonisomorphic K -automorphisms of the same entropy (Ornstein and P. Shields). (b) There is a K -flow that is not Bernoulli, and there are uncountably many nonisomorphic K -flows having the same entropy at time one (Smorodinsky). (c) There exists a K -automorphism not isomorphic to its inverse (Ornstein and Shields). (d) There exists an automorphism that cannot be written as a direct product of a K -automorphism and an automorphism with zero entropy. This example shows that a conjecture made earlier by Pinsker was false (Ornstein). (e) There are weakly isomorphic K -automorphisms that are not isomorphic (Politi and Rudolph).

F. Weak Equivalence and Monotone Equivalence

(1) Weak Equivalence. In order to construct examples of \dagger factors in the theory of von Neumann algebras (\rightarrow 308 Operator Algebras), F. Murray and von Neumann considered various ergodic groups of bimeasurable nonsingular transformations on a finite measure space. In this context a group $\mathcal{G} = \{g\}$ of transformations is called ergodic if $m((g^{-1}(B) \cup B) - (g^{-1}(B) \cap B)) = 0$ for every $g \in \mathcal{G}$ implies either $m(B) = 0$ or $m(X - B) = 0$. A measure μ is said to be invariant under the group $\mathcal{G} = \{g\}$ if μ is invariant under every transformation g in \mathcal{G} . Murray and von Neumann's construction (the so-called **group measure space construction**) gives a type II_1 \dagger factor if the group admits a finite equivalent invariant measure, a type II_∞ factor if it admits an infinite (but σ -finite) equivalent invariant measure, and a type III factor if it has no σ -finite equivalent invariant measure. In this connection we note that Hajian and Y. Itô extended the theorem of Hajian and Kakutani and proved that an arbitrary group $\mathcal{G} = \{g\}$ of nonsingular bimeasurable transformations admits a finite invariant measure if and only if no set of positive measure is weakly wandering under the group \mathcal{G} (a set W is said to be **weakly wandering under a group** \mathcal{G} if there exists an infinite subset $\{g_n | n \in \mathbb{Z}^+\}$ of \mathcal{G} such that $g_n(W) \cap g_k(W) = \emptyset$ for $n \neq k$). Analogously to the terminologies used in the theory of von Neumann algebras, we define ergodic countable group \mathcal{G} (or an ergodic bimeasurable nonsingular transformation φ) to be of **type** II_1 , II_∞ , or III if \mathcal{G} (or φ) has a finite equivalent invariant measure, a σ -finite infinite equivalent invariant measure, or no σ -finite equivalent invariant measure, respectively.

For a bimeasurable nonsingular transformation φ , define the **full group** $[\varphi]$ to be the group of all bimeasurable nonsingular transformations ψ such that, for some $n = n(x)$, $\psi(x)$

$= \varphi^n(x)$ for almost all x . Two transformations φ_1 and φ_2 are said to be **weakly equivalent** if there exists a bimeasurable nonsingular transformation θ such that $\theta[\varphi_1]\theta^{-1} = [\varphi_2]$. H. Dye proved that any pair of type II_1 transformations are weakly equivalent. It easily follows from this that the same is true for type II_∞ transformations. Krieger showed, on the other hand, that among type III transformations there are uncountably many weakly nonequivalent ones. In fact, Krieger [28] introduced an invariant for weak equivalence, called the **ratio set**, by means of which he classified type III transformations into mutually weakly nonequivalent subclasses III_λ for $0 \leq \lambda \leq 1$. Krieger showed further that (a) for $0 < \lambda \leq 1$, every pair of transformations in the class III_λ is weakly equivalent; (b) in the class III_0 there are uncountably many mutually weakly nonequivalent transformations; and (c) two ergodic transformations are weakly equivalent if and only if the corresponding factors constructed via the group measure space construction are \dagger *-isomorphic. An ergodic flow of nonsingular transformations (not necessarily measure-preserving) called the **associated flow** was constructed independently by Krieger [29] and by T. Hamachi, Y. Oka, and M. Osikawa [30], and was shown to give another effective invariant for weak equivalence. Krieger showed that the mapping which assigns to each ergodic transformation its associated flow gives a \dagger bijective mapping between the set of all weak equivalence classes of ergodic type III_0 transformations and the set of all isomorphism classes of aperiodic conservative ergodic flows. In this connection the theorem of U. Krengel and I. Kubo on the representation of ergodic flows, extending the theorem of Ambrose and Kakutani mentioned in Section D, plays a significant role.

A bimeasurable nonsingular transformation θ is called a **normalizer** of another transformation φ if it satisfies $\theta[\varphi] = [\varphi]\theta$. The set of all normalizers $\mathcal{N}(\varphi)$ of a transformation φ forms a group called the **normalizer group** which contains the full group $[\varphi]$ as a subgroup. One can introduce a suitable topology to $\mathcal{N}(\varphi)$ to make it a complete separable metrizable group. Hamachi [31] has shown that, for type III transformations φ , the quotient group $\mathcal{N}(\varphi)/[\varphi]^-$, where $[\varphi]^-$ denotes the closure of $[\varphi]$, is algebraically and topologically isomorphic to the \dagger commutant of the associated flow.

Results obtained by Krieger and others mentioned above were motivated in part by corresponding developments in von Neumann algebra theory due mainly to A. Connes (\rightarrow 308 Operator Algebras). From the recent deep result of Connes on the uniqueness of \dagger approximately finite-dimensional factors of type II_∞

it follows that every approximately finite-dimensional factor with the exception of type III₁ is \ast -isomorphic to a factor constructed from an ergodic transformation via the group measure space construction.

(2) Monotone Equivalence. The theory of weak equivalence discussed above deals with the structure of orbits of a transformation or of groups of transformations. In fact, transformations φ and ψ are weakly equivalent if and only if there exists a bimeasurable nonsingular transformation θ mapping the φ -orbit of almost all x onto the ψ -orbit of $\theta(x)$. A somewhat more stringent notion of equivalence dealing with orbit structure is called **monotone equivalence** or **Kakutani equivalence**. We say that measurable flows $\{\varphi_t\}$ and $\{\bar{\varphi}_t\}$ of measure-preserving transformations on finite measure spaces (X, \mathcal{B}, m) and $(\bar{X}, \bar{\mathcal{B}}, \bar{m})$, respectively, are monotonely equivalent if there exists a bimeasurable measure-preserving transformation θ on X to \bar{X} such that for almost all $x \in X$ and all $t \in \mathbb{R}$, $\theta\varphi_t(x) = \bar{\varphi}_{\tau(t,x)}\theta(x)$, where $\tau(t, x)$ is a monotone increasing function of t . Two S -flows (\rightarrow Section D) built over the same base transformation with different ceiling functions are monotonely equivalent, and it can be shown that monotonely equivalent ergodic flows are isomorphic to S -flows built over the same base transformation. This induces an equivalence relation on transformations, also called monotone equivalence or Kakutani equivalence: Two transformations φ and $\bar{\varphi}$ are monotonely equivalent if they can serve as base transformations of the same (or equivalent) flow. This notion of equivalence was introduced by Kakutani in [32], where he showed that φ and $\bar{\varphi}$ are monotonely equivalent if and only if there are sets $E \subset X$ and $\bar{E} \subset \bar{X}$ such that the induced transformations (\rightarrow Section D) φ_E and $\bar{\varphi}_{\bar{E}}$ are isomorphic. Abramov's formula mentioned in Section E implies that there are at least three classes of transformations (and flows) that are mutually nonequivalent: transformations (flows) of zero, finite, and infinite entropy. Nothing much was done in this equivalence theory for a number of years until in 1975 J. Feldman and E. Satayev independently showed that there are monotonely nonequivalent transformations within each class of zero, positive, and infinite entropy. Since then extensive work has been done by Feldman, Satayev, A. Katok, Ornstein, Weiss, D. Rudolph, M. Ratner, and others, and numerous results have been obtained. For detailed accounts of this development \rightarrow [22, 24, 33, 34]. The main idea employed by Feldman and Satayev was to introduce a new metric called an \bar{f} -metric and then to define a notion corresponding to

V.W.B. of the isomorphism theory described in Section E by using an \bar{f} - instead of a \bar{d} -metric. They then showed that this property, called **loosely Bernoulli (L.B.)** by Feldman and **monotonely very weak Bernoulli (M.V.W.B.)** by Satayev, stays invariant under monotone equivalence, and that there exist transformations with and without this property. An \bar{f} -metric is defined by starting off with an f_N -metric on the space $J^{\{1,2,\dots,N\}}$ instead of the Hamming distance d_N and proceeding in exactly same way as for the definition of \bar{d} , namely, by extending f_N to f_N^ξ and \bar{f}_N and finally to \bar{f} . For α and β in $J^{\{1,2,\dots,N\}}$, $f_N(\alpha, \beta)$ is defined by setting $1 - f_N(\alpha, \beta)$ equal to $1/N$ times the maximal integer n for which there are positive integers $j_1 < j_2 < \dots < j_n$, $k_1 < k_2 < \dots < k_n$ with $\alpha(j_i) = \beta(k_i)$, $i = 1, \dots, n$. Ornstein and Weiss discovered that by substituting \bar{f} for \bar{d} one can develop a theory of monotone equivalence that parallels the isomorphism theory described in Section E. One can define a process (φ, ξ) to be **finitely fixed (F.F.)** by substituting \bar{f} for \bar{d} in the definition of F.D. process, and, as was mentioned earlier, to be L.B. by doing the same in the definition of the V.W.B. process. Ornstein and Weiss proved: (a) If φ and $\bar{\varphi}$ have zero entropy, finite positive, or infinite entropy and if both have F.F. generators, then they are monotonely equivalent. (b) Let φ have an F.F. generator. If $h(\varphi) = 0$, then φ is monotonely equivalent to an irrational rotation of the circle. If $0 < h(\varphi) < \infty$, then φ is monotonely equivalent to a Bernoulli shift of finite entropy. If $h(\varphi) = \infty$, then φ is monotonely equivalent to a Bernoulli shift of infinite entropy. (c) If φ has an F.F. generator, then (φ, ξ) is F.F. for all nontrivial partitions ξ . (d) (φ, ξ) is F.F. if and only if it is L.B. It is known further that within each entropy class there exist uncountably many monotonely nonequivalent transformations. A measurable flow $\{\varphi_t\}$ is called L.B. if it can be represented as an S -flow built over an L.B. transformation. The L.B. flows of zero entropy are those monotonely equivalent to a Kronecker flow on a 2-dimensional torus, and L.B. flows of finite positive (infinite) entropy are those monotonely equivalent to the Bernoulli flow of finite (infinite) entropy. The direct product of an L.B. flow and a Bernoulli flow is L.B., while it was shown by M. Ratner that the horocycle flow (\rightarrow Section G) is L.B. but its direct product with itself is not.

G. Classical Dynamical Systems

By a **classical dynamical system** we mean a \ast diffeomorphism or a flow generated by a smooth \ast vector field on some \ast differentiable

manifold M^n . Such a system is nonsingular with respect to a measure defined by any \dagger Riemannian metric on M^n . For a fixed Riemannian metric, we call measures **smooth** if they have a smooth density with respect to the measure given by the metric.

(1) Among classical dynamical systems, **geodesic flows** have been investigated most extensively. Let $\mathcal{J}_1(M)$ be the unitary \dagger tangent bundle over the manifold M^n . A point $(x, e) \in \mathcal{J}_1(M)$ defines a unique \dagger geodesic through x in the direction of e . The geodesic flow on $\mathcal{J}_1(M)$ is the flow defined by $\varphi_t(x, e) = (x_t, e_t)$, where x_t is the point in M^n reached from x after time t under a motion with unit speed along the geodesic determined by (x, e) , and e_t is the unit vector at x_t tangent to the geodesic. The classical \dagger Liouville theorem in this context implies that the measure on $\mathcal{J}_1(M)$ that is the product of the measure on M^n induced by the metric and the Lebesgue measure on the $(n-1)$ -dimensional sphere gives a smooth invariant measure for the geodesic flow. A wide class of systems arising from mechanics can be described as geodesic flows.

Hopf and G. Hedlund proved that if the manifold M^n is compact and has constant negative curvature, then the geodesic flow is strongly mixing. Later, by using the theory of group \dagger representations, I. Gel'fand and S. Fomin proved that the spectrum of a geodesic flow on a compact manifold of constant negative curvature is Lebesgue, and is even countable Lebesgue in the case where the manifold is of dimension 2. F. Mautner and later L. Auslander, L. Green, and F. Hahn extended this algebraic method to flows obtained under the action of some one-parameter subgroup of a \dagger Lie group acting on its \dagger homogeneous space and obtained extensive results for the case of \dagger nilpotent and some \dagger solvable Lie groups [35].

(2) The flow on an n -dimensional torus defined by

$$\begin{aligned} \varphi_t(x_1, x_2, \dots, x_n) \\ = (x_1 + \omega_1 t, x_2 + \omega_2 t, \dots, x_n + \omega_n t) \end{aligned}$$

is called a **translational flow** or a **Kronecker flow**. The numbers $\omega_1, \omega_2, \dots, \omega_n$ are called **frequencies**. Every orbit of $\{\varphi_t\}$ is dense in the torus if and only if the frequencies are linearly independent over \mathbf{Z} . The motion under a translational flow with independent frequencies is called a **quasiperiodic** motion. A translation flow for a quasiperiodic motion has discrete spectrum.

(3) Sinai obtained a useful criterion for a classical dynamical system to be a K -system. Let M^n be compact, and suppose that $\{\varphi_t\}$ is a flow on M^n defined by a smooth vector field and preserving some smooth measure μ . A

one-parameter group $\{\psi_t\}$ of transformations of the space M^n , given by a vector field, is called the flow **transversal** to the flow $\{\varphi_t\}$ if (i) the decomposition of the space M^n into the trajectories of the flow $\{\psi_t\}$ is invariant under $\{\varphi_t\}$; (ii) the limit $\lim_{s \rightarrow 0} \lim_{t \rightarrow 0} (W_s(t, x) - t)/ts = \alpha(x)$ exists for the function $W_s(t, x)$, which is defined to be the time length of the segment $\{\varphi_s \psi_u(x) | 0 \leq u \leq t\}$ of the trajectory of the flow $\{\psi_t\}$. Sinai's fundamental theorem states that if a flow $\{\varphi_t\}$ is ergodic and has a transversal ergodic flow $\{\psi_t\}$ for which $\int \alpha(x) d\mu < 0$, then $\{\varphi_t\}$ is a K -flow. If $\alpha(x) < 0$, then we can even drop the assumption that $\{\varphi_t\}$ is ergodic. If $\int \alpha(x) d\mu > 0$, the theorem holds for the flow $\{\varphi_{-t}\}$.

A geodesic flow on a 2-dimensional manifold of constant negative curvature always has a transversal flow, called a **horocycle flow**. The ergodicity of a horocycle flow was proved by Hedlund. It follows from Sinai's fundamental theorem, therefore, that a geodesic flow on a surface of constant negative curvature is a K -flow. Sinai proved even more: A geodesic flow on any surface of negative curvature is a K -flow. There is an extension of the notion of transversal flow to higher dimensions, called **transversal field**. Using this notion Sinai proved that a geodesic flow on a manifold (of any dimension) of constant negative curvature is a K -flow. Finally, Ornstein and Weiss established that geodesic flows on compact manifolds of negative curvature are Bernoulli.

(4) D. Anosov considered a class of flows and diffeomorphisms satisfying a condition that characterizes unstable motions such as geodesic flows on a manifold of negative curvature. They are now called **Anosov flows** (or **Y-flows**) and **Anosov diffeomorphisms** (or **Y-diffeomorphisms**) (\rightarrow 126 Dynamical Systems). Anosov proved that if an Anosov flow has a smooth invariant measure, then it is ergodic and either it has a continuous nonconstant eigenfunction or it is a K -flow. Anosov diffeomorphisms with smooth invariant measures are K -automorphisms. Sinai constructed for \dagger transitive Anosov diffeomorphisms (and for Anosov flows) a special partition of the underlying manifold M called a **Markov partition** having desirable properties. The importance of such partitions lies in the fact that they enable one to represent such diffeomorphisms as Markov shifts. Starting with a measure invariant for such a Markov shift having the maximal entropy (\rightarrow Section H), Sinai constructed a Gibbs measure (\rightarrow Section C), which turned out to be unique in this case and gave rise to a natural invariant measure μ for the corresponding diffeomorphism. He proved further that the diffeomorphism φ considered as an

automorphism of the measure space (M, \mathcal{B}, μ) is a K -automorphism. It was shown subsequently by R. Azencott that φ is in fact Bernoulli with respect to μ . Sinai investigated further the uniqueness of the invariant measure attaining the maximal entropy for transitive Anosov diffeomorphisms, and he was able to show among other things that the set of transitive C^∞ -Anosov diffeomorphisms that do not have an invariant measure absolutely continuous with respect to the measure induced by the Riemannian metric on M contains an open dense subset. The methods employed by Sinai in these investigations were extended further by D. Ruelle, Bowen, and others. Bowen, in particular, was able to construct Markov partitions for a wider class of diffeomorphisms, namely, those satisfying the so-called **axiom A** introduced earlier by S. Smale, and characterized the Gibbs measures for diffeomorphisms in this class by means of the **variational principle** (\rightarrow Section H). For a more detailed account of the results of Sinai and Bowen \rightarrow [13, 14].

(5) An important example of a system that is neither an Anosov nor an "axiom A" system because of nonsmoothness has been studied by Sinai: the simplest mechanical model due to Boltzmann and Gibbs of an ideal gas, which is described as a system generated by tiny rigid spherical pellets moving inside a rectangular box and colliding elastically. Sinai succeeded in proving that this is a K -system, thereby giving an affirmative answer to the classical question of the ergodicity of the basic model of statistical mechanics [38]. It was shown subsequently by M. Aizenman, S. Goldstein, and J. Lebowitz that this system is Bernoulli. Sinai also investigated another nonsmooth system of classical importance: the system describing the motion of a billiard ball on a square table with a finite number of convex obstacles. He showed that such a system, is a K -system. Ornstein and G. Gallavotti then showed that Sinai's methods in fact show that the system is Bernoulli. Sinai's methods were extended further by L. Bunimovich and I. Kubo to study properties of billiard systems in more complicated domains. For more detailed account of these matters \rightarrow [24]. The results obtained by Sinai and Bowen for Anosov and "axiom A" systems were extended to even wider class of systems by A. Stepin, R. Sacksteder, M. Brin, and Ya. Pesin (\rightarrow [24]).

H. Miscellany

(1) An arbitrary aperiodic automorphism can be approximated by periodic automorphisms. More precisely, if φ is an aperiodic automor-

phism, then for any positive integer n and $\delta > 0$, there exists a periodic automorphism ψ of period n such that $m(\{x \mid \varphi(x) \neq \psi(x)\}) < (1/n) + \delta$ (theorem of Halmos and Rokhlin). The question as to how quickly this approximation can be carried out has been investigated in detail by Katok, Stepin, and others. The rate of this approximation was shown to have a close relationship with the entropy and spectral properties of the automorphism φ . By utilizing this relationship, various examples of automorphisms with specified spectral properties have been constructed [16].

(2) When an arbitrary automorphism φ is given on a Lebesgue measure space (X, \mathcal{B}, m) , there exists a unique decomposition $\xi = \{A_\lambda\}$ of X such that (i) each A_λ is invariant under φ and (ii) except for a negligible set (in a specified sense) of A_λ , each A_λ is turned (in a natural manner) into a Lebesgue measure space, and the restriction of φ to A_λ is an ergodic automorphism. This decomposition is called the **ergodic decomposition** of X with respect to φ . There is a corresponding decomposition with respect to a flow. A formula also exists that enables us to compute the entropy $h(\varphi)$ in terms of the entropies of ergodic components of φ .

(3) Let X be a compact metric space and $\varphi: X \rightarrow X$ a homeomorphism. N. Krylov and N. Bogolyubov showed that there always exists on X a Borel probability measure μ that is invariant under φ . Let \mathcal{P} be the collection of all Borel probability measures on X , and let \mathcal{P}_φ be the subset of \mathcal{P} consisting of those invariant under φ . Then \mathcal{P} and \mathcal{P}_φ are both convex sets compact with respect to the weak* topology. If \mathcal{E}_φ is the set of all extreme points in \mathcal{P}_φ , then by the Kreĭn-Milman theorem \mathcal{E}_φ is not empty. A measure μ in \mathcal{P}_φ belongs to \mathcal{E}_φ if and only if φ is ergodic with respect to μ . When the set \mathcal{E}_φ consists of a single element, φ is called **uniquely ergodic**; φ is called **minimal** if for every point $x \in X$, the orbit of x under $\varphi = \text{orb}_\varphi(x) = \{\varphi^n(x) \mid n \in \mathbb{Z}\}$ is dense in X . φ is called **strictly ergodic** if it is both minimal and uniquely ergodic. A theorem of J. Oxtoby states that φ is strictly ergodic if and only if for every continuous real-valued function f on X , the sequence of averages $(\sum_{k=0}^{n-1} f(\varphi^k(x)))/n$ converges uniformly to a constant $M(f)$. There are homeomorphisms that are minimal or uniquely ergodic but not strictly ergodic.

(4) R. Jewett proved that any weakly mixing measure-preserving transformation on a Lebesgue space is spatially isomorphic to a strictly ergodic transformation. Krieger extended the result by showing that if the entropy of an ergodic transformation φ is finite, then φ is spatially isomorphic to a strictly ergodic transformation. Similar results were

obtained for flows by K. Jacobs, M. Denker, and E. Eberlein (\rightarrow [39]).

(5) A topological analog of the notion of entropy, called **topological entropy**, was introduced by Adler, A. Konheim, and J. McAndrew. This is defined as follows: For every open covering \mathcal{A} of a compact topological space X , let $N(\mathcal{A})$ be the number of sets in the minimal subcovering of \mathcal{A} . For open coverings \mathcal{A} and \mathcal{B} , let $\mathcal{A} \vee \mathcal{B}$ be the open covering $\{A \cap B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$. For any open covering \mathcal{A} and a continuous mapping φ on X , the limit $\lim_{n \rightarrow \infty} (\log N(\mathcal{A} \vee \varphi^{-1} \mathcal{A} \vee \dots \vee \varphi^{-(n-1)} \mathcal{A})) / n = h_{\text{top}}(\varphi, \mathcal{A})$ exists. Topological entropy $h_{\text{top}}(\varphi)$ of the continuous transformation φ is now defined by $h_{\text{top}}(\varphi) = \sup \{h_{\text{top}}(\varphi, \mathcal{A}) \mid \mathcal{A} \text{ an open covering of } X\}$.

L. Goodwyn showed that $h_{\text{top}}(\varphi) \geq h_{\mu}(\varphi)$ for any φ -invariant probability measure μ , where $h_{\mu}(\varphi)$ is the measure-theoretic entropy of φ regarded as a μ -preserving transformation. T. Goodman went further and succeeded in proving that $h_{\text{top}}(\varphi) = \sup \{h_{\mu}(\varphi) \mid \mu \text{ a } \varphi\text{-invariant probability measure}\}$. In connection with these results there is interest in the question of the existence and uniqueness of an invariant measure for φ with **maximal entropy**, i.e., a φ -invariant measure μ for which $h_{\mu}(\varphi) = h_{\text{top}}(\varphi)$. Such a measure does not always exist, and even if it does it may not be unique. The notions of topological entropy and of measures with maximal entropy are generalized by Ruelle in the following way. For an open covering \mathcal{A} , a continuous mapping φ , and a real-valued continuous function g on X , let $Z_n(\mathcal{A}, \varphi, g)$ be equal to $\inf \{ \sum_{O \in \Gamma} \sup_{x \in O} \exp \sum_{k=0}^{n-1} g(\varphi^k(x)) \}$, where the inf is taken over all subcoverings Γ of the covering $\mathcal{A} \vee \varphi^{-1} \mathcal{A} \vee \dots \vee \varphi^{-(n-1)} \mathcal{A}$. Then a finite limit $P(\mathcal{A}, \varphi, g) = \lim_{n \rightarrow \infty} 1/n \log Z_n(\mathcal{A}, \varphi, g)$ exists, and the quantity $P(\varphi, g) = \sup \{P(\mathcal{A}, \varphi, g) \mid \mathcal{A} \text{ an open covering of } X\}$ is called **topological pressure**. When $g=0$, $P(\varphi, g)$ reduces to $h_{\text{top}}(\varphi)$. Ruelle proved for expansive mappings φ that $P(\varphi, g) = \sup \{h_{\mu}(\varphi) + \int g d\mu \mid \mu \text{ is a } \varphi\text{-invariant probability measure}\}$. This assertion is called the **variational principle for the topological pressure**. The variation principle was proved for general continuous mappings φ by P. Walters. If a φ -invariant measure μ satisfies $P(\varphi, g) = h_{\mu}(\varphi) + \int g d\mu$, then μ is called an **equilibrium state** for g with respect to φ . It is known that for expansive mappings φ every continuous function g on X has an equilibrium state.

(6) Application of ergodic theory to problems in analytic number theory has been made by several authors. Ergodic or mixing properties of particular measure-preserving transformations that arise in connection with various problems in number theory have been ex-

ploited to give answers to these problems.

New and more striking applications of ideas of ergodic theory to different types of questions in number theory have been started by Yu. Linnik, Fürstenberg, W. Veech, T. Kamae, and others (\rightarrow [40, 41]).

7. Most of the results discussed in this article dealt with the action of a cyclic group of transformations or of one-parameter flow. There are significant extensions of many of these results to different types of group actions. For recent developments \rightarrow [22].

References

- [1] E. Hopf, *Ergodentheorie*, Springer, 1937.
- [2] P. R. Halmos, *Lectures on ergodic theory*, Publ. Math. Soc. Japan, 1956.
- [3] P. Billingsley, *Ergodic theory and information*, Wiley, 1965.
- [4] K. Jacobs, *Neuere Methoden und Ergebnisse der Ergodentheorie*, Springer, 1960.
- [5] N. Dunford and J. T. Schwartz, *Linear operators I*, Interscience, 1958.
- [6] K. Yosida, *Functional analysis*, Springer and Kinokuniya, second edition, 1968.
- [7] U. Krengel, Recent progress on ergodic theorems, Soc. Math. France, Astérisque, 50 (1977), 151–192.
- [8] J. F. C. Kingman, Subadditive ergodic theory, Ann. Probability, 1 (1973), 883–909.
- [9] D. Ruelle, Ergodic theory of differentiable dynamical systems, Publ. Math. Inst. HES, 50 (1979), 27–58.
- [10] N. A. Friedman, Introduction to ergodic theory, Van Nostrand, 1969.
- [11] H. Fürstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, J. Analyse Math., 31 (1977), 204–256.
- [12] S. R. Foguel, *Ergodic theory of Markov processes*, Van Nostrand, 1969.
- [13] Ya. G. Sinai, Gibbs measures in ergodic theory, Russian Math. Surveys, 27 (1972), 21–69. (Original in Russian, 1972.)
- [14] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture notes in math. 470, Springer, 1975.
- [15] G. Maruyama, The harmonic analysis of stationary stochastic processes, Mem. Fac. Sci. Kyushu Univ., (A) 4 (1949), 45–106.
- [16] A. B. Katok and A. M. Stepin, Approximations in ergodic theory, Russian Math. Surveys, 22 (5) (1967), 77–102. (Original in Russian, 1967.)
- [17] W. Krieger, On entropy and generators of measure-preserving transformations, Trans. Amer. Math. Soc., 149 (1970), 453–464.
- [18] V. A. Rokhlin (Rohlin), New progress in

the theory of transformations with invariant measure, *Russian Math. Surveys*, 15 (4) (1960), 1–22. (Original in Russian, 1960.)

[19] V. A. Rokhlin, Lectures on the entropy theory of measure-preserving transformations, *Russian Math. Surveys*, 22 (5) (1967), 1–52. (Original in Russian, 1967.)

[20] W. Parry, Entropy and generators in ergodic theory, Benjamin, 1969.

[21] D. S. Ornstein, Ergodic theory, randomness, and dynamical systems, Yale Univ. Press, 1974.

[22] D. S. Ornstein, A survey of some recent results on ergodic theory, *Math. Assoc. Amer. Studies in Math.*, 18 (1978), 229–262.

[23] P. Shields, The theory of Bernoulli shifts, *Chicago Lectures in Math.*, Univ. of Chicago Press, 1973.

[24] A. B. Katok, Ya. G. Sinai, and A. M. Stepin, Theory of dynamical systems and general transformation groups with invariant measure, *J. Soviet Math.*, 7 (1977), 974–1065. (Original in Russian, 1975.)

[25] J. Moser, E. Phillips, and S. Varadhan, Ergodic theory, a seminar, NYU Lecture Notes, 1975.

[26] A. M. Vershik and S. A. Yuzvinskii, Dynamical systems with invariant measure, *Progress in Math.* VIII, Plenum, 1970, 151–215. (Original in Russian, 1967.)

[27] H. A. Dye, On groups of measure-preserving transformations I, II, *Amer. J. Math.*, 81 (1959), 119–159; 85 (1963), 551–576.

[28] L. Sucheston (ed.), Contributions to ergodic theory and probability, Lecture notes in math. 160, Springer, 1970.

[29] W. Krieger, On ergodic flows and the isomorphism of factors, *Math. Ann.*, 223 (1976), 19–70.

[30] T. Hamachi, Y. Oka, and M. Osikawa, Flows associated with ergodic non-singular transformation groups, *Publ. Res. Inst. Math. Sci.*, 11 (1975), 31–50.

[31] T. Hamachi, The normalizer group of an ergodic automorphism of type III and the commutant of an ergodic flow, *J. Functional Anal.*, 40 (1981), 387–403.

[32] S. Kakutani, Induced measure preserving transformations, *Proc. Imp. Acad. Tokyo*, 19 (1943), 635–641.

[33] A. B. Katok, Monotone equivalence in ergodic theory, *Math. USSR-Izv.*, 11 (1977), 99–146. (Original in Russian, 1977.)

[34] E. A. Satayev, An invariant of monotone equivalence determining the quotients of automorphisms monotonely equivalent to a Bernoulli shift, *Math. USSR-Izv.*, 11 (1977), 147–169. (Original in Russian, 1977.)

[35] L. Auslander, L. Green, F. Hahn, et al., Flows on homogeneous spaces, *Ann. Math. Studies*, Princeton Univ. Press, 1963.

[36] V. I. Arnold and A. Avez, Ergodic problems of classical mechanics, Benjamin, 1968.

[37] Ya. G. Sinai, Introduction to ergodic theory, *Math. Notes*, Princeton Univ. Press, 1976.

[38] Ya. G. Sinai, On the foundations of the ergodic hypothesis for a dynamical system of statistical mechanics, *Sov. Math. Dokl.*, 4 (1963), 1818–1822. (Original in Russian, 1963.)

[39] M. Denker, C. Grillenberger, K. Sigmund, Ergodic theory on compact spaces, Lecture notes in math. 527, Springer, 1976.

[40] Yu. V. Linnik, Ergodic properties of algebraic fields, Springer, 1968.

[41] W. A. Veech, Topological dynamics, *Bull. Amer. Math. Soc.*, 83 (1977), 775–830.

137 (VI.18) Erlangen Program

When F. Klein succeeded K. G. C. von Staudt as professor at the Philosophical Faculty of Erlangen University in 1872, he gave an inauguration lecture entitled “Comparative Consideration of Recent Geometric Researches,” which later appeared as an article [1]. In it he developed a penetrating idea, now called the Erlangen program, in which he utilized group-theoretic concepts to unify various kinds of geometries that until that time had been considered separately.

The concept of transformation is not new; it was, however, not until the 18th century that the concept of transformation groups was recognized as useful. The theory of invariants of linear groups and the †Galois theory of algebraic equations attracted attention in the 19th century. In the same century, †projective geometry made remarkable progress, for example, when A. Cayley and E. Laguerre discovered that metrical properties of Euclidean and †non-Euclidean geometries can be interpreted in the language of projective geometry. Cayley proclaimed, “All geometry is projective geometry.” After learning geometry under J. Plücker, Klein made the acquaintance of S. Lie. Both men understood the importance of the group concept in mathematics. Lie studied the theory of †continuous transformation groups, and Klein studied discontinuous transformation groups from a geometric standpoint. Klein was thus led to the idea of the Erlangen program, which provided a bird’s-eye view of geometry.

Klein’s idea can be summarized as follows: A space S is a given set with some geometric structure. Let a transformation group G of S be given. A subset of S , called a **figure**, may

have various kinds of properties. The study of the properties that are left invariant under all transformations belonging to G is called the **geometry of the space S subordinate to the group G** . Let this geometry be denoted by (S, G) . Two figures of S are said to be **congruent** in (S, G) if one of them is mapped to the other by a transformation of G . The geometry (S, G) is actually the theory of invariants of S under G , with the term **invariants** to be understood in a wider sense; it means both invariant quantities and invariant properties or relations.

Replacing G in (S, G) by a subgroup G' of G , we obtain another geometry (S, G') . A series of subgroups of G gives rise to a series of geometries. For instance, let A be a figure of S . The elements of G leaving A invariant form a subgroup $\hat{G}(A)$ of G that operates on $S' = S - A$. We thus obtain a geometry $(S', \hat{G}(A))$ in which A is called an **absolute figure**. In this way, many geometries are obtained from projective geometry. Klein gave numerous examples. It is noteworthy that he mentioned even the groups of \dagger rational and \dagger homeomorphic transformations.

Klein's idea not only synthesized the geometries known at that time, but also became a guiding principle for the development of new geometries.

In 1854 G. F. B. Riemann published his epochmaking idea of **Riemannian geometry**. This geometry has a metric, but in general lacks congruence transformations (isometries). Thus Riemannian geometry is a geometry that is not included in the framework of the Erlangen program. The importance of Riemannian geometry was acknowledged when it was used by A. Einstein in 1916 as a foundation of his general theory of relativity. H. Weyl, O. Veblen, and J. A. Schouten discovered geometries that are generalizations of affine, projective, and \dagger conformal geometries in the same way as Riemannian geometry is a generalization of Euclidean geometry. It became necessary to establish a theory that reconciled the ideas of Klein and Riemann; E. Cartan succeeded in this by introducing the notion of \dagger connection (\rightarrow 80 Connections). However, the Erlangen program, which gave an insight into the essential character of classical geometries, still maintains its role as one of the guiding principles of geometry.

References

[1] F. Klein, *Vergleichende Betrachtungen über neuere geometrische Forschungen*, Math. Ann., 43 (1893), 63–100 (Gesammelte mathematische Abhandlungen, Springer, 1921, vol. 1, 460–497).

- [2] G. Fano, *Kontinuierliche geometrische Gruppen*, Enzykl. Math., Leipzig, 1907–1910, Geometrie III, pt. 1, AB4b, 289–388.
 [3] F. Klein, *Vorlesungen über nicht-Euklidische Geometrie*, Springer, 1928.
 [4] F. Klein, *Vorlesungen über höhere Geometrie*, Springer, 1926.
 [5] E. Cartan, *Les récentes généralisations de la notion d'espace*, Bull. Sci. Math., 48 (1924), 294–320.
 [6] E. Cartan, *La théorie des groupes et les recherches récentes de géométrie différentielle*, Enseignement Math., 24 (1925), 1–18; Proc. Internat. Congr. Math., Toronto, 1 (1928), 85–94.

138 (XV.3) Error Analysis

A. General Remarks

The data obtained by observations or measurements in astronomy, geodesy, and other sciences do not usually give exact values of the quantities in question. The **error** is the difference between the approximation and the exact value. The **theory of errors** originated from systematic work with data accompanied by errors, and the statistical treatment of experimental data was the main concern in the beginning stages (\rightarrow 397 Statistical Data Analysis). However, due to the recent development of high-speed computers it has become possible to carry out computations on a tremendously large scale, and the detailed analysis of errors has become an absolute necessity in modern numerical computation. Hence the analysis of errors in relation to numerical computation has become the center of research in error theory.

B. Errors

One rarely makes a mistake in counting a small number of things; therefore the exact value of the count can be determined. On the other hand, the exact value in decimals is never obtainable for a continuous quantity, say length, no matter how fine measurements are made, and a large or small stochastic error is thus inevitable in measuring a continuous quantity. A discrete finite quantity is a **digital quantity**, and a continuous quantity is an **analog quantity**. The natures of these two quantities are quite different. The values of a digital quantity are distributed on some discrete set, while the values of an analog quantity are distributed with a continuous

probability. Thus there is the possibility of error even for treatment of digital quantities, although checking the results for these quantities is easy. It is preferable to regard digital quantities as being analog quantities if the possible values are densely distributed.

On the other hand, when an appropriate analog computer is not available, analog quantities receive treatment similar to digital quantities. They are expressed as x times some unit, and x is expanded in the decimal or binary systems. An approximation to such an expansion is obtained by rounding off a numeral at some place, the position depending on the capacity for computation by available methods. There are two ways of rounding off numbers, the **fixed point method** and the **floating point method**. The former specifies the place of digits where the rounding off is made, and the latter essentially specifies the number of significant digits.

Classification of errors. (1) **Errors of input data** are the errors included in input data themselves. Such input-data errors include the errors that occur when we represent constants such as $1/3$, $\sqrt{2}$, π by finite decimals. (2) **Truncation errors** occur in approximate expressions for the computation formulas under consideration. (3) **Roundoff errors** occur in taking some finite number of digits from the earlier digits in the numerical value at each step. If the computation of an infinite number of digits were actually possible, no errors of this type would appear. Recently, it has been considered more preferable to call this "computational error."

The difference between fixed point and floating point rounding off is that the former is better suited for operations of addition and subtraction and the latter is better for multiplication and division. In fixed point rounding off, if a number is multiplied many times by numbers less than 1, a so-called **underflow** may occur, and many digits may disappear; a great deal of information can thus be lost. In computation for scientific research that involves frequent multiplication and division, floating point rounding off is preferable. It should be noted that rounding off for addition and subtraction may also cause a critical loss of information. This phenomenon is called **canceling** digits. For instance, in the subtraction $7.6325071 - 7.6318425 = 0.0006646$, where the subtrahend and minuend share several early significant digits, the difference loses those digits. Thus, relative errors may be magnified tremendously. By taking a large number of significant digits, such a situation may be avoided to some extent. So-called **high-precision computation** shows its effectiveness in such cases. Similarly, when a small number b is

added to a large number a , the result may be just a and the information of b could be lost completely. This kind of **loss of information** can often cause serious trouble.

C. Methods of Error Analysis

In order to analyze the **propagation of errors**, let us assume that all numbers are carried to infinitely many digits so that no roundoff error occurs. Suppose that we are to evaluate the function $y = f(x_1, \dots, x_n)$ when x_1, x_2, \dots, x_n are assigned. Let η be the truncation error of an approximate expression. If an input error δ_i for x_i exists, then the corresponding error for y is

$$\eta + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \delta_i.$$

Moreover, suppose that at the final step we round off to get a result with a finite number of digits, by which an error ε is introduced. Then the final error δ for y is

$$\delta = \varepsilon + \eta + \sum_{i=1}^n \frac{\partial f}{\partial x_i} \delta_i.$$

This procedure is performed for each step needed in the computation. If $y = f(x_1, \dots, x_n)$ is a specified step, then the input error δ_i for that step involves all the errors arising before that step, i.e., δ_i is an **accumulated error**. **Forward analysis** is a method to estimate the total accumulated error from the initial input data. It is usually quite difficult to obtain precise estimates by means of this method. In contrast, J. H. Wilkinson proposed the following **backward analysis**. Here, the computational value y is considered as the exact result for the modified initial data $\tilde{x}_1, \dots, \tilde{x}_n$, say $y = f(\tilde{x}_1, \dots, \tilde{x}_n)$, and the estimates for $|x_i - \tilde{x}_i|$ are given. For example, in the binary floating point arithmetic of u bits, we always have the relation:

$$\begin{aligned} &\text{computational value of } a \pm b \\ &= a(1 + \delta) \pm b(1 + \varepsilon), \end{aligned}$$

with $|\delta|, |\varepsilon| \leq 2^{-u}$, even when cancellation or loss of information occurs. Wilkinson has made a deep investigation of error analysis for linear computation, algebraic equations, and eigenvalue problems by means of backward analysis [4, 5].

D. Proliferation of Errors

The phenomena usually called "accumulation of errors" should more appropriately be called the **proliferation** of errors, where the algorithm itself includes a particular mechanism to increase a small error indefinitely. An example is the recursion formula for [†]Bessel functions

stated as

$$J_{n+1}(x) = (2n/x)J_n(x) - J_{n-1}(x).$$

It is customary to compute $J_n(x)$ by this formula, starting with the values of $J_0(x)$ and $J_1(x)$, with x given. By putting $J_{n-1}(x) = y_n$, $J_n(x) = z_n$, the recursion formula can be regarded as a linear transformation of the point $P_n(y_n, z_n)$ in a plane into another point $P_{n+1}(y_{n+1}, z_{n+1})$, where

$$y_{n+1} = z_n, \quad z_{n+1} = -y_n + (2n/x)z_n.$$

The †eigenvalues λ_1, λ_2 of this †difference equation satisfy the following: As long as $n < |x|$, we have $|\lambda_1| = |\lambda_2| = 1$, while if $n > |x|$, λ_1 is greater than 1 and increases rapidly as n tends to infinity. Consequently, even the slightest discrepancy in the position of P_n gives rise to a greatly magnified error in the result [3].

Many studies have been made of the propagation of errors and of instability phenomena in the numerical solution of ordinary differential equations (\rightarrow 303 Numerical Solution of Ordinary Differential Equations; [2]).

References

- [1] J. von Neumann and H. H. Goldstine, Numerical inverting of matrices of higher order, Bull. Amer. Math. Soc., 53 (1947), 1021–1099.
- [2] P. Henrici, Discrete variable methods in ordinary differential equations, Wiley, 1962.
- [3] T. Uno, The problem of error propagation (in Japanese), Sûgaku, 15 (1963), 30–40.
- [4] J. H. Wilkinson, Rounding errors in algebraic processes, Prentice-Hall, 1963.
- [5] J. H. Wilkinson, Algebraic eigenvalue problem, Clarendon Press, 1965.
- [6] L. B. Rall (ed.), Error in digital computation I, II, Wiley, 1965, 1966.

139 (VI.3) Euclidean Geometry

A. History

Attempts to construct axiomatically the geometry of ordinary 3-dimensional space were undertaken by the ancient Greeks, culminating in Euclid's *Elements* (\rightarrow 187 Greek Mathematics). The **fifth postulate** of Euclid's *Elements* requires that two straight lines in a plane that meet a third line, as shown in Fig. 1, in angles α, β whose sum is less than 180° , have a common point. In the *Elements*, two straight lines in a plane without a common point are said to be **parallel**. It can be proved from other

axioms in the *Elements* that if $\alpha + \beta = 180^\circ$, the two lines l and l' in Fig. 2 are parallel. Hence given a line l and a point P not lying on l ,

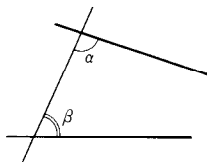


Fig. 1

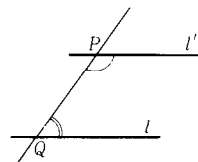


Fig. 2

there exists a line l' passing through P that is parallel to l . The fifth postulate ensures the uniqueness of the parallel l' passing through the given point P . For this reason, the fifth postulate is also called the **axiom of parallels**. Utilizing this axiom, we can prove the well-known theorems on parallel lines, the sum of interior angles of triangles, etc. The axiom plays an important role in the proof of the Pythagorean theorem in the *Elements*. The axiom is also called **Euclid's axiom**.

However, Euclid states this axiom in a quite complicated form, and unlike his other axioms, it cannot be verified within a bounded region of the space.

Many mathematicians tried in vain to deduce it from other axioms. Finally the axiom was shown to be independent of other axioms in the *Elements* by the invention of non-Euclidean geometry in the 19th century (\rightarrow 285 Non-Euclidean Geometry).

The term *Euclidean geometry* is used in contrast to *non-Euclidean geometry* to refer to the geometry based on Euclid's axiom of parallels as well as on other axioms explicit or implicit in Euclid's *Elements*. It was in the 19th century that a complete system of Euclidean geometry was explicitly formulated (\rightarrow 155 Foundations of Geometry). From the standpoint of present-day mathematics, it would be natural to define first the group of motions by the axiom of free mobility due to H. Helmholtz (\rightarrow Section B) and then, following F. Klein, to define Euclidean geometry as the study of properties of spaces that are invariant under the groups of these motions (\rightarrow 137 Erlangen Program).

B. Group of Motions

Let P be an †ordered field and A^n the n -dimensional †affine space over P . Let B^r be an r -dimensional affine subspace of A^n , B^{r-1} an $(r-1)$ -dimensional subspace of B^r , B^{r-2} an $(r-2)$ -dimensional subspace of B^{r-1} , etc. In the sequence of subspaces B^r, B^{r-1}, \dots, B^0 , each $B^k - B^{k-1}$ consists of two †half-spaces ($k = r, r-1, \dots, 1$). Let H^k be one of these half-

spaces. Then the sequence of half-spaces H^r, H^{r-1}, \dots, H^1 is called an r -dimensional **flag**, denoted by \mathfrak{H}^r ($n \geq r \geq 1$), and B^r and H^r are called the **principal space** and the **principal half-space** of \mathfrak{H}^r , respectively. If f is a †proper affine transformation of A^n , $f(H^r), f(H^{r-1}), \dots, f(H^1)$ form an r -dimensional flag \mathfrak{H}^r . We write $f(\mathfrak{H}^r) = \mathfrak{H}^r$.

Let \mathfrak{U}^n be the group of all proper affine transformations of A^n . The subgroup \mathfrak{B}^n of \mathfrak{U}^n with the following two properties is called the **group of motions**, and any element of \mathfrak{B}^n is called a **motion** (or **congruent transformation**).
(1) Let r be an integer between 1 and n , and let $\mathfrak{H}^r, \mathfrak{K}^r$ be any two r -dimensional flags. Then there exists an element f of \mathfrak{B}^n that carries \mathfrak{H}^r to \mathfrak{K}^r : $f(\mathfrak{H}^r) = \mathfrak{K}^r$. (2) Let f, g be two elements of \mathfrak{B}^n with $f(\mathfrak{H}^r) = \mathfrak{K}^r, g(\mathfrak{H}^r) = \mathfrak{K}^r$, and let p be any point on the principal space of \mathfrak{H}^r . Then $f(p) = g(p)$, that is, f, g have the same “effect” on the principal space. In particular, when $r = n$, then $f = g$. That \mathfrak{U}^n possesses a subgroup \mathfrak{B}^n with properties (1) and (2) is called the **axiom of free mobility**.

When $n = 1$, it is easy to see that the elements of \mathfrak{B}^1 are only those elements f of \mathfrak{U}^1 that can be expressed in the form $f(x) = \pm x + a$ ($a \in P$). When $n \geq 2$, P must satisfy the following condition in order that a subgroup \mathfrak{B}^n with properties (1) and (2) exists in \mathfrak{U}^n : If $a, b \in P$, then P contains an element x such that $x^2 = a^2 + b^2$. When this condition is satisfied, the ordered field P is called a **Pythagorean field**. Every †real closed field (e.g., the field \mathbf{R} of real numbers) is Pythagorean. If \mathfrak{B}^n exists, its uniqueness is assured by (1) and (2). Furthermore, if P contains a square root of every positive element (this condition is satisfied, for example, by \mathbf{R}), then conditions (1) and (2) are reducible to the case $r = n$ only, i.e., conditions (1) and (2) for other values of r follow from (1) and (2) with $r = n$. Hereafter, we assume the existence of \mathfrak{B}^n .

Suppose that we have $A^n \supset B^r \supset B^k$ ($n \geq r \geq k \geq 0$), and let \mathfrak{H}^r be a flag with the principal space B^r : $\mathfrak{H}^r = (H^r, \dots, H^k, \dots, H^1)$. Let \mathfrak{K}^r be another flag with the same principal space B^r : $\mathfrak{K}^r = (K^r, \dots, K^k, \dots, K^1)$, where we suppose that $H^j = K^j$ for $k \geq j \geq 1$, whereas for $r \geq i \geq k + 1$, we suppose that H^i and K^i are different half-spaces on B^i divided by B^{i-1} . The flag \mathfrak{K}^r is denoted by \mathfrak{H}_k^r . An element f of \mathfrak{B}^n with $f(\mathfrak{H}^r) = \mathfrak{H}_k^r$ is called a **symmetry** (or **reflection**) of B^r with respect to B^k . It leaves every point on B^k invariant, and its effect on B^r is determined only by B^k independently of the choice of half-spaces in \mathfrak{H}^r and \mathfrak{K}^r (subject to the conditions mentioned above). In particular, the symmetry of A^n with respect to a point $A^0 = p$ is called a **central symmetry** with respect to the **center** p ; and the symmetry of A^n

with respect to a hyperplane $A^{n-1} = h$ is called a **hyperplanar symmetry**. They are uniquely determined by p and h , respectively, and are denoted by S_p and $S(h)$, respectively. If $H(A^n)$ is the set of all hyperplanes of A^n , then \mathfrak{B}^n is generated by $\{S(h) | h \in H(A^n)\}$. Furthermore, if p, q are two points of A^n , the composite $S_p S_q$ is a parallel translation by $2 \cdot \overline{pq}$ (Fig. 3). The parallel translations generate a normal subgroup \mathfrak{T}^n of \mathfrak{B}^n . For $p, q \in A^n$, the element of \mathfrak{T}^n that carries p to q is denoted by τ_{pq} . The

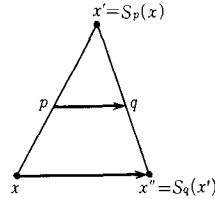


Fig. 3

subgroup of \mathfrak{B}^n that leaves a point p of A^n invariant is denoted by O_p^n . Obviously, we have $O_q^n = \tau_{pq} O_p^n \tau_{pq}^{-1}$. Thus all the O_p^n (for $p \in A^n$) are isomorphic. We call O_p^n the **orthogonal group** around p and any element of O_p^n an **orthogonal transformation** around p . More generally, any element of \mathfrak{B}^n that leaves a subspace A^k of A^n invariant is called an **orthogonal transformation around the subspace A^k** .

An element of \mathfrak{B}^n that preserves the **orientation** of A^n , i.e., is represented by a †proper affinity with a positive determinant, is called a **proper motion**. Proper motions form a subgroup \mathfrak{B}_0^n of \mathfrak{B}^n . Rotations are, by definition, orthogonal transformations belonging to \mathfrak{B}_0^n . Sometimes \mathfrak{B}_0^n is called the group of motions; then \mathfrak{B}^n is called the group of motions in the wider sense. In this article, however, we shall continue to use the terminology introduced above.

The study of the properties of A^n invariant under \mathfrak{B}^n is **n -dimensional Euclidean geometry**. Since $\mathfrak{U}^n \supset \mathfrak{B}^n$, every proposition in affine geometry (\rightarrow 7 Affine Geometry) can be considered a proposition in Euclidean geometry, but there are many propositions that are proper to Euclidean geometry. Sometimes the subgroup of \mathfrak{U}^n generated by \mathfrak{B}^n and the **homotheties** of A^n , i.e., elements of \mathfrak{U}^n represented by †scalar matrices, is called the **group of motions in the wider sense**, and the study of properties of A^n invariant under this group is called **n -dimensional Euclidean geometry in the wider sense**.

C. Length of Segments

Two figures F, F' in A^n are said to be **congruent** if there exists an $f \in \mathfrak{B}^n$ such that $f(F) = F'$.

Then we write $F \equiv F'$. The congruence relation is an \dagger equivalence relation. Let $s = \overline{pq}$, $s' = \overline{p'q'}$ be two \dagger segments in A^n . We say that s, s' have equal **length** when $s \equiv s'$. Length is an attribute of the equivalence class of segments. The length of s is denoted by $|s|$. All segments of the form \overline{pp} are congruent, and we define $|\overline{pp}| = 0$. If we are given a length and a \dagger half-line starting from a point p , we can find a unique point q on it such that $|\overline{pq}| =$ the given length (Fig. 4). Let r be a point on the extension of \overline{pq} . The length $|\overline{pr}|$ is then uniquely determined by $|pq|$ and $|qr|$. It is defined as the

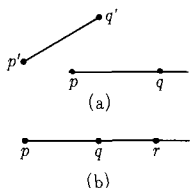


Fig. 4

sum of the lengths: $|\overline{pr}| = |\overline{pq}| + |\overline{qr}|$. With respect to the addition thus defined, lengths of segments in A^n form a \dagger commutative semigroup with the cancellation law, which can be extended to an \dagger Abelian group M with 0 as the identity element (\rightarrow 190 Groups P).

Let $|s| \neq 0$ and $|s'|$ be any length. On a half-line starting from p , we can find points q, r with $|\overline{pq}| = |s|$, $|\overline{pr}| = |s'|$. Then the element $pr/pq = \lambda \in P$ (\rightarrow 7 Affine Geometry) is a positive element of P uniquely determined by $|s|$ and $|s'|$. We call λ the **measure** of $|s'|$ with the **unit** $|s|$ and denote it by $|s'| : |s|$. If P is \dagger Archimedean, λ can be represented by a real number (\rightarrow 149 Fields N). We have $(|s'| + |s''|) : |s| = (|s'| : |s|) + (|s''| : |s|)$, $(|s''| : |s'|)(|s'| : |s''|) = |s''| : |s'|$ (if $|s'| \neq 0$). Thus the mapping $|s'| \rightarrow |s'| : |s|$ sends the additive semigroup of lengths to that of the positive elements of P . This is actually an isomorphism, which can be extended to an isomorphism of M onto the additive group of the field P .

Let $\varphi(X, X', \dots, X^a)$ be a \dagger homogeneous rational function of a finite number of variables X, X', \dots, X^a . If a relation $\varphi(\lambda, \lambda', \dots, \lambda^a) = 0$ holds for $\lambda = |s| : |s_0|$, $\lambda' = |s'| : |s_0|$, \dots , $\lambda^a = |s^a| : |s_0|$, where $|s_0|$ is a length $\neq 0$, then $\varphi(\lambda_1, \lambda'_1, \dots, \lambda_1^a) = 0$ holds also for $\lambda_1 = |s| : |s_1|$, $\lambda'_1 = |s'| : |s_1|$, \dots , $\lambda_1^a = |s^a| : |s_1|$, where $|s_1|$ is any other length $\neq 0$. Hence, in this case, the expression $\varphi(|s|, |s'|, \dots, |s^a|) = 0$ is meaningful.

D. Angles and Their Measure

An **angle** $\angle AOB$ is a figure constituted by two half-lines OA, OB starting from the same point O but belonging to different straight lines. The point O is called the **vertex**, and the two half-

lines OA, OB are called the **sides** of $\angle AOB$ (Fig. 5). Two congruent angles are said to have the same **measure**, denoted by $|\angle AOB|$ or sometimes simply by α . Let $\mathfrak{H}^2 = (H^2, H^1)$ (H^1 = half-line QR) be a given 2-dimensional flag and the given measure of an angle. Then we can find a unique half-line QP in the given half-plane H^2 such that $|\angle PQR| = \alpha$ (Fig. 6). The angle $\angle PQR$ is said to “belong” to \mathfrak{H}^2 . Let K^2 be the half-plane separated by the line $P \cup Q$ containing the half-line QR . Then $H^2 \cap K^2$ is called the **interior** of the angle

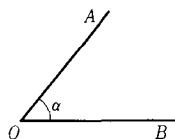


Fig. 5

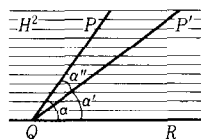


Fig. 6

$\alpha' + \alpha'' = \alpha$.

$\angle PQR$. Let $\angle P'QR$ be another angle belonging to \mathfrak{H}^2 . If the interior of the latter angle is a subset of the interior of $\angle PQR$ and $QR \neq QP'$, then $|\angle PQR|$ is said to be greater than $|\angle P'QR|$, and we write $|\angle PQR| > |\angle P'QR|$. Actually, it can be shown that $>$ is a relation between the measures of $\angle PQR$ and $\angle P'QR$, and that the set of measures of angles forms a \dagger linearly ordered set with respect to the relation \geq defined in the obvious way. When $|\angle PQR| > |\angle P'QR|$, we write $|\angle PQR| = |\angle PQP'| + |\angle P'QR|$. Actually, these are relations between measures of angles. Furthermore, if the measure α of an angle is given, the set of measures of angles $\leq \alpha$ forms a linearly ordered set order-isomorphic to a segment and satisfying: (i) If $\beta < \gamma$ then there exists a δ such that $\beta + \delta = \gamma$; (ii) $\beta + \delta = \delta + \beta$; $(\beta_1 + \beta_2) + \beta_3 = \beta_1 + (\beta_2 + \beta_3)$ if all these sums exist; and (iii) $\beta_1 + \delta = \beta_2 + \delta$ implies $\beta_1 = \beta_2$. When P is Archimedean, these properties imply that the measure of angles $\leq |\angle PQR|$ can be represented by positive real numbers $\leq k$ (k is any given positive number) such that the relations of ordering an addition are preserved.

This one-to-one correspondence between the measures of angles and a subset S of the interval $(0, k]$ of real numbers can be extended to a correspondence between the measures of **general angles** and the subset of \mathbf{R} obtained from S . When $P = \mathbf{R}$, then we have $S = (0, k]$, and any real number appears as a measure of a general angle. We can choose $\angle PQR$ and the positive number k arbitrarily, but it is customary to choose them as follows. Suppose we are given an angle $\angle AOB$. Let the extensions of the half-lines OA and OB in the opposite directions be OA' and OB' , respectively. The angles $\angle AOB$ and $\angle A'OB$ are called **supplementary angles** of each other, and so are

$\angle AOB$ and $\angle A'OB'$. The angles $\angle AOB$ and $\angle A'OB'$ are called **vertical angles** to each other (Fig. 7). Any angle is congruent to its vertical angle, and an angle that is congruent to its supplementary angle has a fixed measure.

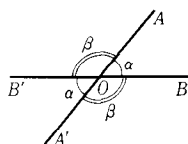


Fig. 7

Such an angle (or its measure) is called a **right angle**. In the description of the measurement of angles, we usually consider the case where the special angle $\angle PQR$ is a right angle, and we set $k = \pi/2$. (The existence and uniqueness of the right angle can be proved.) An angle that is greater (smaller) than a right angle is called an **obtuse (acute) angle**. A general angle whose measure is twice (four times) a right angle is called a **straight angle (perigon)**. Sometimes we choose as the "unit angle" $1/90$ of a right angle, which is called a **degree** (hence a right angle = 90 degrees, denoted by 90°); $1/60$ of a degree is called a **minute** ($1^\circ = 60$ minutes, denoted by $60'$), and $1/60$ of a minute is called a **second** ($1' = 60$ seconds, denoted by $60''$). If, as usual, we put the right angle equal to $\pi/2$, then the unit angle is $(2/\pi)(\text{right angle})$. This is called a **radian**, and $1 \text{ radian} = 180^\circ/\pi = 57^\circ 17' 44.806 \dots'' \approx 57.3^\circ$.

If a straight line m intersects two straight lines l, l' , eight angles $\alpha, \beta, \gamma, \delta, \alpha', \beta', \gamma', \delta'$ appear, as in (Fig. 8). In this figure, α and α', β and β', γ and $\gamma',$ and δ and δ' are called **corresponding angles**, while α and γ', β and δ', γ and $\alpha',$ and δ and β' are called **alternate angles** to each other. When l and l' are parallel, each of these angles is congruent to its corresponding or alternate angle.

The **Pythagorean theorem** asserts that if a triangle $\triangle ABC$ is given for which $\angle ABC$ is a right angle (Fig. 9), then $|AB|^2 + |BC|^2 = |CA|^2$ (which makes sense since $X^2 + Y^2 - Z^2$ is a homogeneous polynomial).

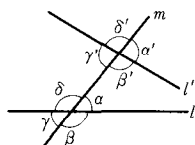


Fig. 8

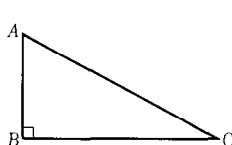


Fig. 9

E. Rectangular Coordinates

When two straight lines l, m intersect, two pairs of vertical angles appear. If one of these angles is a right angle, then all are. Then we

say that l and m are **orthogonal** (or **perpendicular**) to each other, and write $l \perp m$. Let l be a line and A^r an r -dimensional subspace of A^n ($1 \leq r \leq n-1$) intersecting l at a point $O = A^r \cap l$. If l is orthogonal to all lines on A^r passing through O , then l is said to be orthogonal to A^r , and we write $l \perp A^r$ (Fig. 10). If A^{n-1} is any hyperplane in A^n , then there exists a unique line l through a given point P of A^n that is orthogonal to A^{n-1} ; this l is called the **perpendicular** to A^{n-1} through P , and the

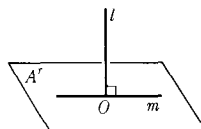


Fig. 10

intersection $l \cap A^{n-1}$ is called the **foot of the perpendicular** through P . When A^{n-1} is given, the mapping from A^n to A^{n-1} assigning to every point P of A^n the foot of the perpendicular through P is called the **orthogonal projection** from A^n to A^{n-1} .

Let $\mathfrak{H}^n = (H^n, H^{n-1}, \dots, H^1)$ be an n -dimensional flag of A^n and O the initial point of the half-line H^1 . Then we can find a point E_i in H^i ($i = 1, 2, \dots, n$) such that $O \cup E_i \perp O \cup E_j$ ($i \neq j, i, j = 1, 2, \dots, n$). Moreover, if $|e|$ is any unit of length, then E_i can be chosen uniquely so that $|OE_i| = |e|$ ($i = 1, 2, \dots, n$). Then O, E_1, \dots, E_n are independent points in A^n , and we have $A^n = O \cup E_1 \cup \dots \cup E_n$. Thus we have a frame $\Sigma = (O; E_1, \dots, E_n)$ of A^n with O as origin and the E_i as unit points. Such a frame is called an **orthogonal frame**. A coordinate system with this frame, called an **orthogonal coordinate system adapted to \mathfrak{H}^n** , is uniquely determined by \mathfrak{H}^n . A motion is characterized as an affinity sending one orthogonal frame onto another or onto itself.

Utilizing an orthogonal coordinate system, the lengths of segments and the measures of angles can be expressed simply. Let (x_1, \dots, x_n) be the coordinates of X with respect to such a coordinate system. Then the length of the segment $|OX|$ (with $|e|$ as unit) is equal to $(\sum_{i=1}^n x_i^2)^{1/2}$, and when Y is another point, with coordinates (y_1, \dots, y_n) , $O \neq X, O \neq Y$, then we have

$$\cos |\angle XOY| = \frac{\sum_{i=1}^n x_i y_i}{(\sum_{i=1}^n x_i^2)^{1/2} (\sum_{i=1}^n y_i^2)^{1/2}}.$$

In particular, we have $O \cup X \perp O \cup Y$ if and only if $\sum_{i=1}^n x_i y_i = 0$.

We may write $\mathbf{x} = \overrightarrow{OX}$ for the location vector of X . Then the affinity $A\mathbf{x} + \mathbf{b}$ is a motion if and only if A is an orthogonal matrix. Thus the inner product (\mathbf{x}, \mathbf{y}) is invariant

under motions; it therefore has meaning in Euclidean geometry. If we put $|\mathbf{x}| = (\mathbf{x}, \mathbf{x})^{1/2}$, then the right-hand sides of the formulas for $|OX|$ and $\cos \angle XOY$ can be written as $|\mathbf{x}|$ and $(\mathbf{x}, \mathbf{y})/(|\mathbf{x}| \cdot |\mathbf{y}|)$. More generally, we have $|XY| = |\mathbf{y} - \mathbf{x}|$. This is the **Euclidean distance** (or simply **distance**) between X and Y . Then A^n becomes a \dagger metric space with this distance, i.e., a Euclidean space. Historically, the notion of metric spaces was introduced in generalizing Euclidean spaces (\rightarrow 273 Metric Spaces).

F. Area and Volume

The subset I^n of A^n consisting of points (x_1, \dots, x_n) with respect to an orthogonal coordinate system, with $0 \leq x_i \leq 1$, $i = 1, \dots, n$, is called an n -dimensional **unit cube**. A function m that assigns to \dagger polyhedra in the wider sense P, Q, \dots in A^n nonnegative real numbers $m(P), m(Q), \dots$ is called an n -dimensional **volume** if it satisfies the following four conditions: (1) $m(\emptyset) = 0$. (2) $m(P \cup Q) + m(P \cap Q) = m(P) + m(Q)$. (3) If P is sent to Q by a translation, then $m(P) = m(Q)$. (4) $m(I^n) = 1$. It has been proved that such a function is unique and has the property that $P \equiv Q$ implies $m(P) = m(Q)$. Thus the concept of volume can be defined in the framework of Euclidean geometry. More generally, if the affinity $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ sends P onto Q , then $m(Q) = cm(P)$, where c is the absolute value of the determinant $|A|$. If P is covered by a finite number of hyperplanes, then $m(P) = 0$, and if P is a \dagger parallelootope with n independent edges $\mathbf{a}_1, \dots, \mathbf{a}_n$, then $m(P) = \text{abs}|\mathbf{a}_1, \dots, \mathbf{a}_n|$, where $|\mathbf{a}_1, \dots, \mathbf{a}_n|$ is the determinant of the $n \times n$ matrix with \mathbf{a}_i as the i th column vector, and $\text{abs } x$ is the absolute value of the real number x . If P is an n -simplex whose vertices have location vectors $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$, then we have

$$m(P) = \frac{1}{n!} \text{abs} \begin{vmatrix} 1 & 1 & \dots & 1 \\ \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{vmatrix}.$$

The volume of any polyhedron can be obtained by dividing it into n -simplexes and summing their volumes. If P is an r -dimensional polyhedron in A^n , then the r -dimensional volume of P is obtained by dividing P into r -simplexes and summing their r -dimensional volumes (in the respective r -dimensional Euclidean spaces containing them). In particular, when $r = 1$, we speak of **length** (e.g., the length of a broken line), and when $r = 2$, of **area**. If V is the r -dimensional volume of an r -dimensional parallelootope with r edges $\mathbf{a}_1, \dots, \mathbf{a}_r$, we have the formula

$$V^2 = \begin{vmatrix} (\mathbf{a}_1, \mathbf{a}_1) & \dots & (\mathbf{a}_1, \mathbf{a}_r) \\ \vdots & \ddots & \vdots \\ (\mathbf{a}_r, \mathbf{a}_1) & \dots & (\mathbf{a}_r, \mathbf{a}_r) \end{vmatrix}.$$

The notion of measure of point sets other than polyhedra is a generalization of the notion of volume of polyhedra (\rightarrow 270 Measure Theory).

G. Orthonormalization

Let O, A_1, \dots, A_n be $n+1$ \dagger independent points in A^n . Then the n vectors $\overrightarrow{OA_i} = \mathbf{a}_i$, $i = 1, \dots, n$, are independent. The points O, A_1, \dots, A_n determine an n -dimensional flag \mathfrak{H}^n of A^n as follows. Let H_1 be the half-line OA_1 , H_2 be the half-plane on the plane $O \cup A_1 \cup A_2$ separated by the line $O \cup A_1$ in which A_2 lies, \dots , H_n be the half-space on A^n separated by the hyperplane $O \cup A_1 \cup \dots \cup A_{n-1}$ in which A_n lies. Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be the unit vectors of the rectangular coordinate system adapted to \mathfrak{H}^n . Suppose further that we are given a rectangular coordinate system. Then $\mathbf{b}_1, \dots, \mathbf{b}_n$ can be obtained from $\mathbf{a}_1, \dots, \mathbf{a}_n$ by the following procedure, called **orthonormalization** (E. Schmidt): First put $\mathbf{b}_1 = \mathbf{a}_1/|\mathbf{a}_1|$, so that $|\mathbf{b}_1| = 1$. Then $\mathbf{c}_2 = \mathbf{a}_2 - (\mathbf{a}_2, \mathbf{b}_1)\mathbf{b}_1$ satisfies $(\mathbf{b}_1, \mathbf{c}_2) = 0$, $\mathbf{c}_2 \neq 0$. Put $\mathbf{b}_2 = \mathbf{c}_2/|\mathbf{c}_2|$. Then we have $(\mathbf{b}_1, \mathbf{b}_2) = 0$, $|\mathbf{b}_2| = 1$. When $\mathbf{b}_1, \dots, \mathbf{b}_{i-1}$ are obtained in this way, so that $(\mathbf{b}_j, \mathbf{b}_k) = \delta_{jk}$ for $1 \leq j, k \leq i-1$, then $\mathbf{c}_i = \mathbf{a}_i - (\mathbf{a}_i, \mathbf{b}_1)\mathbf{b}_1 - \dots - (\mathbf{a}_i, \mathbf{b}_{i-1})\mathbf{b}_{i-1}$ satisfies $(\mathbf{b}_j, \mathbf{c}_i) = 0$, $\mathbf{c}_i \neq 0$. Hence $\mathbf{b}_i = \mathbf{c}_i/|\mathbf{c}_i|$ added to $\mathbf{b}_1, \dots, \mathbf{b}_{i-1}$ retains the property $(\mathbf{b}_j, \mathbf{b}_k) = \delta_{jk}$ for $1 \leq j, k \leq i$, and this procedure can be continued to $i = n$.

Two vectors \mathbf{u}, \mathbf{v} are called **orthogonal** (denoted $\mathbf{u} \perp \mathbf{v}$) if $(\mathbf{u}, \mathbf{v}) = 0$, and \mathbf{u} is called **normalized** when $|\mathbf{u}| = 1$. Thus any two of the vectors $\mathbf{b}_1, \dots, \mathbf{b}_n$ are orthogonal, and each of them is normalized. Between given vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ and $\mathbf{b}_1, \dots, \mathbf{b}_n$ we have the relation $\{\mathbf{a}_1, \dots, \mathbf{a}_i\}$ ($=$ the linear space generated by $\mathbf{a}_1, \dots, \mathbf{a}_i$) $= \{\mathbf{b}_1, \dots, \mathbf{b}_i\}$, $i = 1, \dots, n$.

Let $\mathfrak{M}_1, \mathfrak{M}_2$ be two subspaces of the linear space \mathfrak{M} of the vectors of the Euclidean space A^n . If any element of \mathfrak{M}_1 is orthogonal to any element of \mathfrak{M}_2 , then \mathfrak{M}_1 and \mathfrak{M}_2 are called **orthogonal** and written $\mathfrak{M}_1 \perp \mathfrak{M}_2$. For any proper subspace \mathfrak{M}_1 of \mathfrak{M} , it can be shown by the method of orthonormalization that there exists a unique proper subspace \mathfrak{M}_2 of \mathfrak{M} such that $\mathfrak{M} = \mathfrak{M}_1 \cup \mathfrak{M}_2$, $\mathfrak{M}_1 \perp \mathfrak{M}_2$. Such a subspace \mathfrak{M}_2 is called the **orthocomplement** of \mathfrak{M}_1 (with respect to \mathfrak{M}). Then $\mathfrak{M}_1 \cap \mathfrak{M}_2 = \{0\}$ follows, and hence $\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2$. Every element \mathbf{a} of \mathfrak{M} is therefore written uniquely in the form $\mathbf{a}_1 + \mathbf{a}_2$, $\mathbf{a}_1 \in \mathfrak{M}_1$, $\mathbf{a}_2 \in \mathfrak{M}_2$; we call \mathbf{a}_1 the \mathfrak{M}_1 -component of \mathbf{a} and \mathbf{a}_2 the **orthogonal component** of \mathbf{a} with respect to \mathfrak{M}_1 . The mapping from \mathfrak{M} to \mathfrak{M}_1 assigning \mathbf{a}_1 to \mathbf{a} is called the **orthogonal projection** from \mathfrak{M} to \mathfrak{M}_1 ; it is a linear and \dagger idempotent mapping.

H. Distance between Subspaces

Since the Euclidean space A^n is a metric space, the distance is defined between any two non-empty subsets of A^n (\rightarrow 273 Metric Spaces).

Let A^r, B^s be two subspaces of dimensions r, s of A^n , and let d be the distance between them. Then it can be shown that there exist points $p \in A^r, q \in B^s$ such that $d = \overline{pq}$, and if $p' \in A^r, q' \in B^s$ are any other points with $d = \overline{p'q'}$, then $\overline{pq} = \overline{p'q'}$. In particular, when $r = 0$ and $s = n - 1$ (i.e., when $A^r = p$ is a point and $B^s = B^{n-1}$ is a hyperplane), the distance d can be obtained as follows: If $(\mathbf{a}, \mathbf{x}) = b$ is an equation of B^{n-1} and \mathbf{p} is the location vector of p , then $d = |(\mathbf{a}, \mathbf{p}) - b|/|\mathbf{a}|$. If $|\mathbf{a}| = 1$ in this equation of B^{n-1} , then d is given simply by $|(\mathbf{a}, \mathbf{p}) - b|$. An equation $(\mathbf{a}, \mathbf{x}) = b$ of a hyperplane is said to be in **Hesse's normal form** if $|\mathbf{a}| = 1$.

I. Spheres and Subspaces

The set of points in a Euclidean space lying at a fixed distance r from a given point is called the **sphere of radius r** with **center** at the given point. If \mathbf{p} is the location vector of the center of this sphere with respect to a given rectangular coordinate system, then the equation of the sphere is $|\mathbf{x} - \mathbf{p}| = r$ or $(\mathbf{x}, \mathbf{x}) - 2(\mathbf{p}, \mathbf{x}) + (\mathbf{p}, \mathbf{p}) - r^2 = 0$. The set of points lying at equal distances from $k + 1$ points with location vectors $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_k$ ($k \geq 1$) is a linear subspace of the space (which may be \emptyset or the entire space). If these points are independent, then the subspace has dimension $n - k$, where n is the dimension of the entire space. In particular, if these points are vertices of an n -dimensional \dagger simplex, then there is a unique sphere passing through them, called the **circumscribing sphere** of the simplex. In this case, the simplex is said to be **inscribed** in the sphere. If $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_n$ are location vectors of the vertices of the simplex, then the equation of the circumscribing sphere of the simplex is given by

$$\begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ \mathbf{p}_0 & \mathbf{p}_1 & \dots & \mathbf{p}_n & \mathbf{x} \\ \mathbf{p}_0^2 & \mathbf{p}_1^2 & \dots & \mathbf{p}_n^2 & \mathbf{x}^2 \end{vmatrix} = 0.$$

When $n = 2$ or 3 , there are many classical results concerning the circumscribing circle of a triangle, the circumscribing sphere of a simplex, and other figures related to a triangle or a simplex.

References

[1] G. D. Birkhoff and R. Beatley, *Basic geometry*, Scott, Foresman, 1941 (Chelsea, third edition, 1959).

[2] E. E. Moise, *Elementary geometry from an advanced standpoint*, Addison-Wesley, 1963.

[3] H. Weyl, *Mathematische Analyse des Raumproblems*, Springer, 1923 (Chelsea, 1960).

[4] F. Klein, *Vorlesungen über höhere Geometrie*, Springer, third edition, 1926 (Chelsea, 1957).

[5] J. Dieudonné, *Algèbre linéaire et géométrie élémentaire*, Hermann, second corrected edition, 1964; English translation, *Linear algebra and geometry*, Hermann, 1969.

[6] G. Choquet, *L'enseignement de la géométrie*, Hermann, 1964.

140 (VI.4)
Euclidean Spaces

A space satisfying the axioms of Euclidean geometry is called a **Euclidean space**. An \dagger affine space having as \dagger standard vector space an n -dimensional Euclidean \dagger inner product space over a real number field \mathbf{R} is an n -dimensional Euclidean space E^n . In an n -dimensional Euclidean space E^n , we fix an \dagger orthogonal frame $\Sigma = (O, E_1, \dots, E_n)$, $e_i = \overline{OE_i}$, $(e_i, e_j) = \delta_{ij}$. The frame Σ determines \dagger rectangular coordinates (x_1, x_2, \dots, x_n) of each point in E^n . We can thus establish a one-to-one correspondence between E^n and $\mathbf{R}^n = \{(x_1, \dots, x_n) | x_i \in \mathbf{R}\}$. In this sense we identify E^n and \mathbf{R}^n and usually call \mathbf{R}^n itself a Euclidean space. The 1-dimensional space \mathbf{R}^1 is a straight line, and the \dagger Cartesian product of n copies of \mathbf{R}^1 is an n -dimensional Euclidean space (or **Cartesian space**). Given points $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ in the Euclidean space \mathbf{R}^n , the \dagger distance $d(x, y)$ between them is given by

$$\sqrt{(y_1 - x_1)^2 + \dots + (y_n - x_n)^2}.$$

Thus the distance $d(x, y)$ supplies \mathbf{R}^n with the structure of a \dagger metric space. We call x_i the i th coordinate of the point x , the point $(0, \dots, 0)$ the **origin** of \mathbf{R}^n , and the set of points $\{x | -\infty < x_i < \infty; x_j = 0, j \neq i\}$ the x_i -**axis** (or i th **co-ordinate axis**). For an integer m such that $-1 \leq m \leq n$, we define m -dimensional \dagger subspaces in \mathbf{R}^n ; a -1 -dimensional subspace is the empty set, a 0 -dimensional subspace is a point, and a 1 -dimensional subspace is a straight line. If we take an orthogonal frame, an m -dimensional subspace is represented as an \mathbf{R}^m (\rightarrow 139 Euclidean Geometry; 7 Affine Geometry).

As a \dagger topological space, \mathbf{R}^n is \dagger locally compact and \dagger connected. A bounded closed set in \mathbf{R}^n is \dagger compact (**Bolzano-Weierstrass theorem**).

Given a point $a = (a_1, \dots, a_n)$ in \mathbf{R}^n and a real positive number r , the subset $\{x | d(x, a) \leq r\}$ of

\mathbf{R}^n is called an n -dimensional **solid sphere**, **solid n -sphere**, **ball**, **n -ball**, **disk**, or **n -disk** with center a and radius r , its \dagger interior $\{x \mid d(x, a) < r\}$ an n -dimensional **open sphere**, **open n -sphere**, **open ball**, **open n -ball**, **open disk**, or **open n -disk**, and its \dagger boundary $\{x \mid d(x, a) = r\}$ an $(n-1)$ -dimensional **sphere** or **$(n-1)$ -sphere**. In particular, a 2-dimensional solid sphere is called a **circular disk**, its interior an **open circle**, and its boundary a **circumference**. A disk or a circumference is sometimes called simply a **circle**.

The family of n -dimensional open spheres with center a gives a base for a neighborhood system of the point a . Suppose that we are given a sphere S and two points x, y on S . The points x, y are called **antipodal points** on the sphere S if there exists a straight line L passing through the center of S such that $S \cap L = \{x, y\}$. The segment (or the length of the segment) whose endpoints are antipodal points is called the **diameter** of the solid sphere (or of the sphere). The notion of \dagger diameter (\rightarrow 273 Metric Spaces) of a solid sphere or of a sphere considered as a subset of the metric space \mathbf{R}^n coincides with the notion of diameter of the corresponding set defined above. When $n \geq 3$, the intersection of a sphere and a 2-dimensional plane passing through the center of the sphere is called a **great circle** of the sphere. For m such that $1 \leq m \leq n$, we consider an m -dimensional solid sphere or an $(m-1)$ -dimensional sphere in an m -dimensional plane \mathbf{R}^m . These spheres are also called m -dimensional solid spheres or $(m-1)$ -dimensional spheres in \mathbf{R}^n .

In particular, the solid sphere of radius 1 having the origin as its center is called the **unit disk**, **unit ball**, or **unit cell**, and its boundary is called the **unit sphere**. (In particular, when we deal with the 2-dimensional space \mathbf{R}^2 , we use the term *circle* instead of *sphere*, as in **unit circle**.) The points $(0, \dots, 0, 1)$ and $(0, \dots, 0, -1)$ are called the **north pole** and **south pole** of the unit sphere, respectively. The $(n-2)$ -dimensional sphere, which is the intersection of the unit sphere and the hyperplane $x_n = 0$, is called the **equator**; the part of the unit sphere that is "above" this hyperplane (i.e., in the half-space $x_n \geq 0$) is called the **northern hemisphere**, and the part that is "below" the hyperplane (i.e., in the half-space $x_n \leq 0$) the **southern hemisphere**.

Let a_i, b_i be real numbers satisfying $a_i < b_i$ ($i = 1, 2, \dots, n$). The subset $\{x \mid a_i < x_i < b_i, i = 1, 2, \dots, n\}$ of \mathbf{R}^n is called an **open interval** of \mathbf{R}^n , and the subset $\{x \mid a_i \leq x_i \leq b_i\}$ a **closed interval**. They are sometimes called **rectangles** (when $n = 2$), **rectangular parallelepipeds**, or **boxes**. An open interval is actually an open set of \mathbf{R}^n , and a closed interval is a closed set. In particular, the closed interval $\{x \mid 0 \leq x_i \leq 1, i = 1, 2, \dots, n\}$ is called the **unit cube** (or **unit n -**

cube) of \mathbf{R}^n . We can take the set of open intervals as base for a neighborhood system of \mathbf{R}^n .

All the \dagger convex closed sets (for example, closed intervals) having interior points in \mathbf{R}^n are homeomorphic to an n -dimensional solid sphere. A topological space I^n that is homeomorphic to an n -dimensional solid sphere is called an n -dimensional (**topological**) **solid sphere**, (**topological**) **n -cell**, or **n -element**. A topological space S^{n-1} homeomorphic to an $(n-1)$ -dimensional sphere is called an $(n-1)$ -dimensional **topological sphere** (or simply $(n-1)$ -dimensional sphere). The spaces I^n and S^{n-1} are \dagger orientable \dagger topological manifolds whose orientations are determined by assigning the generators of the (relative) \dagger homology groups $H_n(I^n, \dot{I}^n)$ and $H_{n-1}(S^{n-1})$, respectively (both are infinite cyclic groups). By means of the \dagger boundary operator $\partial: H_n(I^n, \dot{I}^n) \rightarrow H_{n-1}(S^{n-1})$, the orientation of I^n or S^{n-1} determines that of the other.

References

See references to 139 Euclidean Geometry.

141 (XXI.22) Euler, Leonhard

Leonhard Euler (April 15, 1707–September 18, 1783) was born in Basel, Switzerland. In his mathematical development he was greatly influenced by the Bernoullis (\rightarrow 38 Bernoulli Family). He was invited to the St. Petersburg Academy in 1726 and remained there until 1741, when he was invited to Berlin by Frederick the Great (1712–1786). Euler was active at the Berlin Academy until 1766, when he returned to St. Petersburg. Already having lost the sight of his right eye in 1735, he now became blind in his left eye also. This, however, did not impede his research in any way, and he continued to work actively until his death in St. Petersburg.

Euler was the central figure in the mathematical activities of the 18th century. He was interested in all fields of mathematics, but especially in analysis in the style of \dagger Leibniz, which had been passed down through the Bernoullis and was developed by him into a form that led to the mathematics of the 19th century. Through his work analysis became more easily applicable to the fields of physics and dynamics. He developed calculus further and dealt formally with complex numbers. He also contributed to such fields as \dagger partial differential equations, the theory of \dagger elliptic func-

tions, and the calculus of variations. He contributed much to the progress of algebra and theory of numbers in this period, and also did pioneering work in topology. He had, however, little of the concern for rigorous foundations that characterized the 19th century. He was the most prolific mathematician of all time, and his collected works are still incomplete, though some seventy volumes have already been published.

References

- [1] L. Euler, Opera omnia, ser. 1, vol. 1–29, ser. 2, vol. 1–30, ser. 3, vol. 1–13, Teubner and O. Füssli, 1911–1967.

142 (XV.11) Evaluation of Functions

A. History

By the evaluation of a function $f(x)$ we mean the application of algorithms for obtaining the approximate value of the function. The evaluation methods are classified roughly into two groups: (1) evaluation of functions using approximation formulas, and (2) evaluation using microprogramming techniques. The advent of high-speed computers has brought about drastic changes in the evaluation of functions. Before the introduction of high-speed computers in the 1940s, mathematical tables had played a prominent role. The first issue (published in 1943) of the journal *Mathematical tables and other aids to computation* (MTAC), the predecessor of the journal *Mathematics of computation*, was primarily concerned with tables of mathematical functions. One of the aims of this journal was to facilitate the exchange of information on errors in the tables. The tables, obtained in the past by tedious hand calculation, can now be easily prepared by high-speed computers, and there was a period when more accurate and extensive tables were published one after another. It is ironic, however, that high-speed computers revolutionized numerical analysis and prompted a shift of emphasis in the field, beginning in the 1950s, away from the use of numerical tables and toward exploration of the most efficient methods of approximation of the functions, thus causing a rapid decrease in the need for tables.

Recently, a significant trend in computer design has replaced the conventional logic control section with “stored logic,” or microprogrammed control, stored in high-speed,

nondestructive Read-Only Storage (ROS). For example, the microprogrammed control used for the unified COordinate Rotation DIGital Computer (CORDIC) algorithm is effective in calculating elementary functions because of its simplicity, accuracy, and capability of high-speed execution via parallel processing. It is not clear whether the applications of approximation formulas heretofore in use will be superseded by microprogramming techniques in all kinds of computers. However, the value of the approximation formulas recognized in the 1950s has been declining insofar as elementary functions are concerned. Recently, as a method for the evaluation of functions based on a new viewpoint, some “unrestricted” algorithms have been proposed by Brent [1], which are useful for the computation of elementary and special functions when the required precision is not known in advance or when high accuracy is necessary. It is expected that methods for the evaluation of functions using approximation formulas will be further developed as microprogramming techniques and unrestricted algorithms see wider use.

B. Evaluation of Functions Using Approximation Formulas

Suppose we approximate a function $f(x)$ by a function $g(x)$ using the following class of functions $p_i(x)$ and $q_j(x)$. Let continuous functions $p_1(x), \dots, p_n(x)$ and $q_1(x), \dots, q_m(x)$ defined on a closed interval $[a, b]$ satisfy the following conditions: (i) p_1, \dots, p_n and q_1, \dots, q_m are both linearly independent; (ii) there exist at most a finite number of zeros for $\sum_{j=1}^m b_j q_j(x)$ in $[a, b]$ for any choice of b_1, b_2, \dots, b_m , except for the case $b_1 = b_2 = \dots = b_m = 0$; (iii) there is a continuous function $g(x)$ with a nonzero denominator in $[a, b]$ representable as

$$g(x) = \sum_{i=1}^n a_i p_i(x) \Big/ \sum_{j=1}^m b_j q_j(x), \quad (1)$$

where $a_i, i = 1, \dots, n$, and $b_j, j = 0, 1, \dots, m$, are constants. Then $g(x)$ is called a **generalized rational function** based on a class of functions $\{p_i(x)\}, i = 1, \dots, n$, and $\{q_j(x)\}, j = 1, \dots, m$. If $p_i(x) = x^{i-1}$ and $q_j(x) = x^{j-1}$, (1) is reduced to a rational function. If $m = 1$ and $q_1(x) = 1$, (1) is a linear combination of $p_i(x)$ and is called an approximation of linear type to $f(x)$; and further, if $p_i(x) = x^{i-1}$, (1) is reduced to a polynomial. The crux of the approximation problem lies in the criterion to be used in choosing the approximate constants in (1). There are three methods for choosing them, which lead to three types of approximation of major importance: (i) interpolatory approximations, (ii) least-squares approximations, and (iii)

min-max error approximations (sometimes called **best approximations**). As a major aim of computer approximation of a function is to make the maximum error as small as possible, the third type of approximation has been used for digital computers.

For every continuous function $f(x)$, there always exist min-max error approximations of the form (1). In generating a min-max error approximation in practice, we essentially depend upon the following conditions and theorems. Let a function space F be a d -dimensional linear space. If for $f \in F$ which is not identically zero, there exist at most $(d-1)$ zeros of $f(x)$ in $[a, b]$, then F is called the **unsolvent** space or the **Haar space**.

Let one of the best approximations to a continuous function $f(x)$ be $g(x)$. If the linear space D of all the functions of the form $\{\sum a_i p_i(x) + \sum b_k g(x) q_k(x)\}$ is unsolvent, then the best approximation is unique. For the case where $g(x)$ is an approximation of linear type to $f(x)$, Haar [2] proved the following theorem.

Let $g(x)$ be the best approximation to $f(x)$ and $g(x) = \sum_{i=1}^m a_i p_i(x)$. A necessary and sufficient condition for the best approximation to be always unique is that the m -dimensional linear space generated by the functions p_1, p_2, \dots, p_m be unsolvent. In particular, we have the best unique polynomial approximation (one variable) of $f(x)$ defined on $[a, b]$.

Let the necessary and sufficient condition of the previous theorem be satisfied in a d -dimensional linear space D . A necessary and sufficient condition for $g(x)$ given by (1) to be the best approximation of $f(x)$ is that there exist points $a \leq x_0 < x_1 < \dots < x_{d-1} < x_d \leq b$, called deviation points, such that $(-1)^i(f(x_i) - g(x_i)) = \mu$ ($i=0, 1, \dots, d$).

A great number of algorithms are known by means of which one can calculate the best approximation $g(x)$ of a function $f(x)$ for the given values of n and m and for $p_i(x) = x^{i-1}$ and $g_j(x) = x^{j-1}$ in (1). However, the following three kinds of algorithms are used most frequently for generating min-max approximations on computers: Remes's second algorithm [3], the differential correction algorithm, and Yamauti's folding-up method.

C. Evaluation of Elementary Functions Based on Microprogramming

Chen [6] has given a general algorithm for calculating an elementary function $z = f(x)$ as follows. Let $F(x, y) = yg(x) + h(x)$, where y is some parameter for evaluating $z_0 = f(x_0) = F(x_0, y_0)$. We assume that (x_0, y_0) has been given and that z_0 is unknown. Suppose that a

new point (x_{k+1}, y_{k+1}) is obtained from (x_k, y_k) according to a pair of transformations, $x_{k+1} = \varphi(x_k, y_k)$ and $y_{k+1} = \psi(x_k, y_k)$, keeping the value of $F(x_0, y_0)$ invariant. If x_k is forced to converge to x_ω and if $g(x_\omega) = 1$ and $h(x_\omega) = 0$, then $z_0 = f(x_0) = F(x_0, y_0) = F(x_1, y_1) = \dots = F(x_\omega, y_\omega) = y_\omega$. In this procedure it is necessary to determine $g(x)$, $h(x)$, $\varphi(x)$, and $\psi(x)$ for the given $f(x)$. Most of the elementary functions can be evaluated by Chen's algorithm, which is essentially identical with Specker's Sequential Table Look-up (STL) method based on addition formulas [7], e.g., $\log x = \log(x\alpha) - \log \alpha$ or $e^x = e^{x-\beta} e^\beta$. The iteration equations (2) of CORDIC given later can also be derived by the STL method with complex numbers.

The use of coordinate rotation to evaluate elementary functions is not new. In 1956 and 1959 Volder [8] described the CORDIC for the calculation of trigonometric functions, multiplication, division, and conversion between the binary and the decimal and r -adic number systems. Daggett, in 1959, discussed the use of CORDIC for decimal-binary conversion. It was recognized by Walther [9] in 1971 that these algorithms could be merged into one unified algorithm. Consider coordinate systems parameterized by m in which the radius R and angle A of the vector $P = (x, y)$ are defined as $R = (x^2 + my^2)^{1/2}$ and $A = m^{-1/2} \arctan(m^{1/2} y/x)$. The basis of Walther's algorithm is the coordinate rotation in a linear ($m=0$), circular ($m=1$), or hyperbolic ($m=-1$) coordinate system, depending on which function is calculated. Iteration equations of CORDIC are as follows. The point $(x_{i+1}, y_{i+1}, z_{i+1})$ is obtained from the point (x_i, y_i, z_i) by means of the transformation

$$\begin{aligned} x_{i+1} &= x_i + m y_i \delta_i, \\ y_{i+1} &= y_i - x_i \delta_i, \\ z_{i+1} &= z_i + \alpha_i, \end{aligned} \quad (2)$$

where m is a parameter for the coordinate system, $\alpha_i = m^{-1/2} \arctan(m^{1/2} \delta_i)$, and δ_i is a suitable value, e.g., $\pm 2^{-i}$. The angle A_{i+1} and radius R_{i+1} are $A_{i+1} = A_i - \alpha_i$ and $R_{i+1} = R_i \times K_i \equiv R_i \times (1 + m \delta_i^2)^{1/2}$. After n iterations we find $A_n = A_0 - \alpha$ and $R_n = R_0 \times K$, and then

$$\begin{aligned} x_n &= K \{x_0 \cos(\alpha m^{1/2}) + y_0 m^{1/2} \sin(\alpha m^{1/2})\}, \\ y_n &= K \{y_0 \cos(\alpha m^{1/2}) - x_0 m^{-1/2} \sin(\alpha m^{1/2})\}, \\ z_n &= z_0 + \alpha, \end{aligned}$$

where $\alpha = \sum_{i=0}^{n-1} \alpha_i$ and $K = \prod_{i=0}^{n-1} K_i$. These relations are summarized in Table 1 for $m=1$, $m=0$, and $m=-1$ in the following special cases. (i) The value of A is forced to converge to zero; $y_n \rightarrow 0$. (ii) The value of z is forced to converge to zero; $z_n \rightarrow 0$.

Table 1. Input and Output Functions for CORDIC

Function	m	Input			Quantity to be 0	Output
		x_0	y_0	z_0		
$\sin t$	1	$1/K$	0	t	$z_n \rightarrow 0$	$y_n \rightarrow \sin t$
$\cos t$	1	$1/K$	0	t	$z_n \rightarrow 0$	$x_n \rightarrow \cos t$
$\tan^{-1} t$	1	1	t	0	$y_n \rightarrow 0$	$z_n \rightarrow \tan^{-1} t$
xz	0	x	0	z	$z_n \rightarrow 0$	$y_n \rightarrow xz$
y/x	0	x	y	0	$y_n \rightarrow 0$	$z_n \rightarrow y/x$
$\sinh t$	-1	$1/K_{-1}$	0	t	$z_n \rightarrow 0$	$y_n \rightarrow \sinh t$
$\cosh t$	-1	$1/K_{-1}$	0	t	$z_n \rightarrow 0$	$x_n \rightarrow \cosh t$
$\tanh^{-1} t$	-1	1	t	0	$y_n \rightarrow 0$	$z_n \rightarrow \tanh^{-1} t$

$$K_1 = \prod_{j=0}^{n-1} (1 + \delta_j^2)^{1/2}, \quad K_{-1} = \prod_{j=0}^{n-1} (1 - \delta_j^2)^{1/2}$$

D. Fast Fourier Transform (FFT)

When a function $f(x)$ can be taken to be periodic, it is advantageous to use trigonometric polynomials as least-squares approximating functions. The summations arising in least-squares approximations based on the trigonometric polynomials play an important role in various applications, and a quite efficient algorithm for the evaluation of such sums was developed by Cooley and Tukey [10] in 1965, known as the **fast Fourier transform (FFT)**. Let x_m , ($0 \leq m \leq N-1$) be a set of complex numbers, and consider

$$X_n = (1/N) \sum_{m=0}^{N-1} x_m \exp(-2\pi i m n / N) \quad (0 \leq n \leq N-1). \quad (3)$$

Equation (3) is often called the **discrete Fourier transform (DFT)** of the sequence x_m , it being analogous to the continuous Fourier transform,

$$X_n = (1/T) \int_0^T x(t) \exp(-2\pi i n t / T) dt \quad (0 \leq n \leq N), \quad (4)$$

where $x(t)$ is a periodic function of t with period T . X_n is called the n th Fourier coefficient of a set of N equally spaced samples of size N for the function $x(t)$ ($t = jT/N, 0 \leq j \leq N$). In the same way as for the continuous Fourier transform, the discrete transform can be inverted to yield

$$x_n = \sum_{m=0}^{N-1} X_m \exp(2\pi i m n / N) \quad (0 \leq n \leq N-1).$$

Here x_n is called the coefficient of the inverse Fourier transform, and X_m and x_n thus form a transform pair.

The FFT algorithm for a finite sequence of length N in (3) is based on the fact that the calculation of (3) can be performed in stages by using the direct product decomposition if $N = N_1 N_2$, with N_1 and N_2 relatively prime. For

example, the 2-dimensional Fourier transform coefficient is given by

$$X_{n_1, n_2} = [1/(N_1 N_2)] \times \sum_{m_1=0}^{N_1-1} \sum_{m_2=0}^{N_2-1} \exp(-2\pi i (n_1 m_1 / N_1 + n_2 m_2 / N_2)) x_{m_1, m_2} \quad (0 \leq n_1 \leq N_1-1, 0 \leq n_2 \leq N_2-1).$$

If by an elementary operation we mean one complex multiplication and one complex addition, we can evaluate X_{n_1, n_2} through $(N_1 N_2)^2$ such operations using Horner's scheme. By the direct product decomposition method, however, the $(N_1 N_2)^2$ operations can be reduced to only $N_1 N_2 (N_1 + N_2)$ operations. Because the matrix corresponding to the transformation mentioned above is a direct product of $N_1 \times N_1$ and $N_2 \times N_2$ matrices, we can perform the calculations in two stages: first to obtain ξ_{m_1, n_2} for $0 \leq m_1 \leq N_1-1$ and $0 \leq n_2 \leq N_2-1$ and then to obtain X_{n_1, n_2} for $0 \leq n_1 \leq N_1-1$ and $0 \leq n_2 \leq N_2-1$. We have

$$\xi_{m_1, n_2} = (1/N_2) \sum_{m_2=0}^{N_2-1} \exp(-2\pi i n_2 m_2 / N_2) x_{m_1, m_2}, \\ X_{n_1, n_2} = (1/N_1) \sum_{m_1=0}^{N_1-1} \exp(-2\pi i n_1 m_1 / N_1) \xi_{m_1, n_2}.$$

This direct product decomposition method is well known for 2- or 3-dimensional Fourier transforms. Even for a 1-dimensional Fourier transform of length $N = N_1 N_2$, if N_1 and N_2 are relatively prime, one can use the method of direct product decomposition. Even when N_1 and N_2 are not prime, we can use a method of "pseudo"-direct product decomposition to reduce the number of operations. Namely, we can put $m = m_1 + N_1 m_2$ ($0 \leq m_1 < N_1, 0 \leq m_2 < N_2$); $n = N_2 n_1 + n_2$ ($0 \leq n_1 < N_1, 0 \leq n_2 < N_2$). Then, similarly as before, X_n in (3) can be rewritten as

$$X_n = (1/N_1) \sum_{m_1=0}^{N_1-1} \exp(-2\pi i n_1 m_1 / N_2) \times \exp(-2\pi i m_1 n_2 / (N_1 N_2)) \xi_{m_1, n_2},$$

where

$$\xi_{m_1, n_2} = (1/N_2) \sum_{m_2=0}^{N_2-1} \exp(-2\pi i n_2 m_2 / N_2) x_{m_1 + N_1 m_2}, \quad (5)$$

which is a DFT of length N_2 . If we put

$$\tilde{\xi}_{m_1, n_2} = \exp(-2\pi i n_1 n_2 / (N_1 N_2)) \xi_{m_1, n_2}, \quad (6)$$

then

$$X_n = (1/N_1) \sum_{m_1=0}^{N_1-1} \exp(-2\pi i n_1 m_1 / N_1) \tilde{\xi}_{m_1, n_2}. \quad (7)$$

Formula (7) is nothing but a DFT of length N_1 . Thus an FFT of length $N = N_1 N_2$ can be calculated by decomposing it into three stages as follows: (i) obtain N_1 transforms of length N_2 in (5); (ii) multiply ξ_{m_1, n_2} by $\exp(-2\pi i n_1 n_2 / N)$ in (6) (phase rotation); (iii) calculate N_2 transforms of length N_1 in (7). If either or both of N_1 and N_2 can be factored further so that, e.g., $N = N_1 N_2 = N_1 N_{21} N_{22} = \dots$, an FFT of length N_2 can be calculated similarly by decomposing it, and so on, and in this way one can reduce the total number of operations. This is the principle of FFT pointed out by H. Takahasi. When N is a power of two, it can be shown that the FFT algorithm requires approximately $N \log_2 N/2$ operations.

E. Padé Approximation

Let $f(x) \sim c_0 + c_1 x + c_2 x^2 + \dots$ be a formal power series. For any pair of nonnegative integers (p, q) , we define the (p, q) th **Padé approximation** of $f(x)$ as follows: The Padé approximation is a rational function

$$(a_0 + a_1 x + a_2 x^2 + \dots + a_p x^p) / (b_0 + b_1 x + b_2 x^2 + \dots + b_q x^q)$$

satisfying the condition that all terms in the formal power series

$$(b_0 + b_1 x + \dots + b_q x^q)(c_0 + c_1 x + \dots) - (a_0 + a_1 x + \dots + a_p x^p)$$

should vanish up to the term x^{p+q} . An infinite matrix whose (p, q) th entry is the (p, q) th Padé approximation is called the **Padé table** for $f(x)$. The Padé approximation is uniquely determined, provided that every **Hankel determinant**

$$\begin{vmatrix} c_p & c_{p+1} & \dots & c_{p+q} \\ c_{p+1} & c_{p+2} & \dots & c_{p+q+1} \\ \dots & \dots & \dots & \dots \\ c_{p+q} & c_{p+q+1} & \dots & c_{p+2q} \end{vmatrix}$$

never vanishes.

When we expand $f(x)$ into a continued

fraction of Stieltjes type, say,

$$f(x) \sim \frac{c_0}{1} + \frac{\alpha_1 x}{1} + \frac{\alpha_2 x}{1} + \dots,$$

then its $2p$ th and $(2p+1)$ th approximate fractions are the (p, p) th and $(p+1, p)$ th Padé approximation of $f(x)$, respectively.

References

- [1] R. P. Brent, Unrestricted algorithms for elementary and special functions, IFIP, 1980, 613–619.
- For the theory of approximation,
- [2] A. Haar, Die Minkowskische Geometrie und die Annäherung an stetige Funktionen, Math. Ann., (1918), 294–311.
- [3] A. Ralston, Rational Chebyshev approximation by Remes' algorithms, Numer. Math., 7 (1965), 322–330.
- [4] E. W. Cheney, Introduction to approximation theory, McGraw-Hill, 1966.
- [5] M. J. D. Powell, Approximation theory and methods, Cambridge Univ. Press, 1981.
- For microprogramming techniques,
- [6] T. C. Chen, The automatic computation of exponential, logarithms, ratios and square roots, IBM Res. Rep. RJ 970 (1972), 32.
- [7] W. H. Specker, A class of algorithms for $\ln x$, $\exp x$, $\sin x$, $\cos x$, $\tan^{-1} x$ and $\cot^{-1} x$, IEEE Trans. Elec. Comp., EC-14 (1965), 85–86.
- [8] J. E. Volder, Binary computation algorithms for coordinate rotation and function generation, Convair Rep. IAR-I 148 Aero-electronics group, 1956.
- [9] J. S. Walther, A unified algorithm for elementary functions, Proc. AFIPS, 1971, Spring Joint Comp. Conf., 379–385.
- For fast Fourier transforms,
- [10] J. W. Cooley and J. W. Tukey, An algorithm for the machine calculation of complex Fourier series, Math. Comp., 19 (1965), 297–301.
- [11] E. C. Brigham, The fast Fourier transform, Prentice-Hall, 1974.
- For the Padé approximation,
- [12] H. Padé, Sur la représentation approchée d'une fonction par des fractions rationnelles, Ann. Sci. Ecole Norm. Sup., (3) 9 (1892), 1–92.
- [13] G. A. Baker, Jr., The Padé approximation 1, 2, Encyclopedia of Mathematics and Its Applications, vols. 13, 14, Addison-Wesley, 1981.
- For tables of approximation formulas,
- [14] C. Hastings, Jr., Approximations for digital computers, Princeton Univ. Press, 1955.
- [15] J. F. Hart (ed.), Computer approximation, Wiley, 1968.

143 (XI.18) Extremal Length

A. General Remarks

The notable relation between the lengths of certain families of curves in a plane domain and the area of the domain has long been recognized and utilized in function theory. L. V. Ahlfors and A. Beurling formulated this relation by introducing the notion of extremal length for families of curves [1]. Although there are various definitions of extremal length, they are essentially the same except for one due to J. Hersch [3] and A. Pfluger [2].

The image of an interval under a continuous mapping is called a curve. We say that it is **locally rectifiable** if every †continuous arc of the curve is †rectifiable. Let C be a finite or countable collection of locally rectifiable curves in a plane and ρ , $0 \leq \rho \leq \infty$, be a †Baire function defined in the plane. Represent C in terms of arc length s (\rightarrow 246 Length and Area), and set $\langle C, \rho \rangle = \int_C \rho ds$. For a family Γ of finite or countable collections C , ρ is called **admissible** if $\langle C, \rho \rangle \geq 1$ for every $C \in \Gamma$. If no $C \in \Gamma$ consists of a finite or countable number of points, then $\rho \equiv \infty$ is always admissible for Γ . Call $\inf \{ \int \rho^2 dx dy \}$, where ρ runs over admissible Baire functions, the **module** of Γ , and denote it by $M(\Gamma)$. The reciprocal $\lambda(\Gamma) = 1/M(\Gamma)$ is called the **extremal length** of Γ . If no Baire functions are admissible for Γ , we set $\lambda(\Gamma) = 0$. The extremal length is defined equivalently in two other ways as follows: Let ρ be a nonnegative Baire function, and put $L(\Gamma, \rho) = \inf \{ \langle C, \rho \rangle \mid C \in \Gamma \}$, then $\lambda(\Gamma) = \sup L(\Gamma, \rho)^2 / \int \rho^2 dx dy$, where ρ runs over nonnegative Baire functions. Next, let Φ be the collection of nonnegative Baire functions ρ such that $\int \rho^2 dx dy \leq 1$; then $\lambda(\Gamma) = \sup \{ L(\Gamma, \rho)^2 \mid \rho \in \Phi \}$. We obtain the same value for $\lambda(\Gamma)$ if we require an admissible ρ to be †lower semicontinuous. If ρ is required to be continuous, then the **extremal length defined by Hersch and Pfluger** is obtained. As is shown in example (1) of Section B, there is a case where the two definitions actually differ.

If an admissible ρ yields $M(\Gamma) = \int \rho^2 dx dy$, then $\rho |ds|$ is called an extremal metric. Beurling gave a necessary and sufficient condition for a metric to be extremal [4].

We list four properties of extremal length:

(1) $\lambda(\Gamma_1) \geq \lambda(\Gamma_2)$ if $\Gamma_1 \subset \Gamma_2$. (2) $M(\bigcup_n \Gamma_n) \leq \sum_n M(\Gamma_n)$. (3) Let $\{\Gamma_n\}$ and Γ be given. Suppose that there are mutually disjoint measurable sets $\{E_n\}$ such that each $C_n \in \Gamma_n$ is contained in E_n . If each element of $\bigcup_n \Gamma_n$ contains at least one $C \in \Gamma$, then $M(\Gamma) \geq \sum_n M(\Gamma_n)$, and

hence

$$M\left(\bigcup_n \Gamma_n\right) = \sum_n M(\Gamma_n).$$

If each $C \in \Gamma$ contains at least one $C_n \in \Gamma_n$ for every n , then $\lambda(\Gamma) \geq \sum_n \lambda(\Gamma_n)$. (4) Let f be an analytic function in a domain Ω and $\{C\}$ be given in Ω . Denote by $f(C)$ the image of C by f . Then $\lambda(\{C\}) \leq \lambda(\{f(C)\})$. The equality holds if f is one-to-one. This shows that $\lambda(\{C\})$ is conformally invariant.

B. Extremal Distance

Let Ω be a domain in a plane, $\partial\Omega$ its boundary, and X_1, X_2 sets on $\Omega \cup \partial\Omega$. The extremal length of the family of curves in Ω connecting points of X_1 and points of X_2 is called the **extremal distance** between X_1 and X_2 (relative to Ω) and is denoted by $\lambda_\Omega(X_1, X_2)$.

Example (1). Let $\Omega = \{z \mid |z| < 2\}$, $X_1 = \partial\Omega$, and X_2 be a countable set in $|z| < 1$ such that the set of accumulation points of X_2 coincides with $|z| = 1$. Then $\lambda_\Omega(X_1, X_2) = \infty$, but the extremal distance in the sense of Hersch and Pfluger is equal to $(2\pi)^{-1} \log 2$.

Example (2). In a rectangle with sides a and b , the extremal distance between the sides of length a is b/a .

Example (3). Let Ω be an annulus $r_1 < |z| < r_2$. The extremal distance λ between the two boundary circles of Ω is equal to $(2\pi)^{-1} \log(r_2/r_1)$. The extremal length of the family of curves in Ω homotopic to the boundary circles is equal to $1/\lambda$.

Example (4). Let Ω be a domain in the extended z -plane such that $\infty \in \Omega$. Let $z_0 \in \Omega$ and $\{|z - z_0| = r\} \subset \Omega$, and denote by λ_r the extremal distance between $\{|z - z_0| = r\}$ and a set $X \subset \partial\Omega$ relative to Ω . Then $\lambda_r - (2\pi)^{-1} \log r$ increases with r . We call the limit the **reduced extremal distance** and denote it by $\tilde{\lambda}_\Omega(X, \infty)$. †Robin's constant for †Green's function in Ω with pole at $z = \infty$ is equal to $2\pi \tilde{\lambda}_\Omega(\partial\Omega, \infty)$.

Extremal length is also defined on Riemann surfaces. Some classical conformal invariants can be given in a generalized form in terms of extremal length. The notion of extremal length has applications in various branches of function theory, such as †conformal and †quasi-conformal mappings, the †Phragmén-Lindelöf theorem, the †coefficient problem, the †type problem of Riemann surfaces, and studies of the boundary properties of functions of finite Dirichlet integrals. It is also applied to problems in differential geometry. Extending the notion of extremal length, M. Ohtsuka considered **extremal length with weight**, and B. Fuglede introduced the notion of **generalized**

module in higher-dimensional spaces [8]. These notions have useful properties and applications.

References

- [1] L. V. Ahlfors and A. Beurling, Conformal invariants and function-theoretic null-sets, *Acta Math.*, 83 (1950), 101–129.
- [2] A. Pfluger, Extremallängen und Kapazität, *Comment. Math. Helv.*, 29 (1955), 120–131.
- [3] J. Hersch, Longueurs extrémales et théorie des fonctions, *Comment. Math. Helv.*, 29 (1955), 301–337.
- [4] M. Ohtsuka, Dirichlet problem, extremal length and prime ends, Van Nostrand, 1970.
- [5] Y. Kusunoki, Some classes of Riemann surfaces characterized by the extremal length, *Proc. Japan Acad.*, 32 (1956).
- [6] M. Ohtsuka, On limit of BLD functions along curves, *J. Sci. Hiroshima Univ.*, 28 (1964).
- [7] T. Fujiï'e, Boundary behaviors of Dirichlet functions, *J. Math. Kyoto Univ.*, 10 (1970).
- [8] B. Fuglede, Extremal length and functional completion, *Acta Math.*, 98 (1957), 171–219.
- [9] J. A. Jenkins, Univalent functions and conformal mapping, *Erg. Math.*, Springer, 1958.
- [10] J. Väisälä, Lectures on n -dimensional quasiconformal mappings, Lecture notes in math. 229, Springer, 1971.

F

144 (XXI.23) Fermat, Pierre de

Pierre de Fermat (August 20, 1601–January 12, 1665) was born into a family of leather merchants near Toulouse, France. He became an attorney and in 1631 a member of the Toulouse district assembly. When not engaged in such work, he did research in mathematics, so that he consigned his results only to his correspondence or to unpublished manuscripts. The manuscripts were published posthumously by his son in 1679 and are known as *Varia opera mathematica*. His research into number theory, stimulated by Bachet's (1581–1638) translation of the *Arithmetika* of Diophantus (published in 1621), made Fermat's name immortal and initiated modern number theory. He posed the famous Fermat's Problem, which has yet to be solved (\rightarrow 145 Fermat's Problem). He began analytic geometry by studying the theory of †conic sections of Apollonius, and utilizing this theory he dealt with the notions of tangent lines, maximal (minimal) values of functions, and quadrature, which made him a pioneer in calculus. He also wrote a precursory work in the theory of probability in the course of his correspondence with †Pascal. †Fermat's principle is important in the field of optics, where it is known as the law of least action. Unlike †Descartes, he emphasized the revival rather than the criticism of Greek mathematics.

References

[1] P. Fermat, *Oeuvres I–V and supplement*, P. Tannery and C. Henry (eds.), Gauthier-Villars, 1891–1922.

145 (V.16) Fermat's Problem

The **last theorem of Fermat** (c. 1637) asserts that if n is a natural number greater than 2, then

$$x^n + y^n = z^n \quad (1)$$

has no rational integral solution x, y, z with $xyz \neq 0$. In the case $n=2$, equation (1) has integral solutions called **Pythagorean numbers** (\rightarrow 118 Diophantine Equations). Fermat read a Latin translation of Diophantus' *Arithmetika*, in which the problem of finding all Pythagorean numbers is treated. In his personal copy of that book, Fermat wrote his as-

sertion as a marginal note at the point at which $n=2$ in equation (1) is treated and added the famous words, "I have discovered a truly remarkable proof of this theorem which this margin is too small to contain." It is not known whether Fermat actually had a proof. Fermat's problem asks for a proof or disproof of this conjecture, which itself has not been solved despite centuries of efforts by many mathematicians; but its study has promoted remarkable advances in number theory. In particular, E. E. Kummer's theory of ideal numbers and the development of the theory of †cyclotomic fields were originally conceived in treating Fermat's problem.

In this article, we consider only those integral solutions x, y, z of equation (1) with $xyz \neq 0$ that are relatively prime. We also restrict ourselves to the cases $n=l$ (odd prime) and $n=4$, without loss of generality.

For smaller values of n , the nonsolvability of equation (1) was proved long ago, for $n=3$ by L. Euler (1770), and later again by A. M. Legendre; for $n=4$ by Fermat and Euler; for $n=5$ by P. G. L. Dirichlet and Legendre (1825); and for $n=7$ by G. Lamé (1839). S. Germain and Legendre found some results on more general cases, but the most remarkable result was obtained by Kummer (*J. Reine Angew. Math.*, 40 (1850), *Abh. Akad. Wiss. Berlin* (1857)).

Let l be an odd prime, ζ a primitive l th root of unity, and h the †class number of the cyclotomic field $\mathbf{Q}(\zeta)$. Then the class number h_2 of the real subfield $\mathbf{Q}(\zeta + \zeta^{-1})$ of $\mathbf{Q}(\zeta)$ divides h . We call $h_1 = h/h_2$ and h_2 the †first and †second factors of h , respectively.

(1) If l is †regular, that is, if $(h, l) = 1$, then $x^l + y^l = z^l$ has no solution (Kummer, 1850).

There are infinitely many irregular primes [3]; those under 100 are 37, 59, and 67. There are 7,128 regular primes and 4,605 irregular primes between 3 and 125,000. It is not yet known whether there are infinitely many regular prime numbers, although the beginning part of the sequence of natural numbers contains a larger number of these than the number of irregular prime numbers. The condition $(l, h) = 1$ is equivalent to saying that the numerators of †Bernoulli numbers B_{2m} ($m = 1, 2, \dots, (l-3)/2$) are not divisible by l (Kummer, 1850).

Kummer obtained a result on irregular primes (1857) which was improved later as follows. Note that if l is not regular then h_1 is divisible by l (Kummer, 1850) (\rightarrow 14 Algebraic Number Fields).

(2) If $(h_2, l) = 1$ and the numerators of Bernoulli numbers B_{2m} ($m = 1, 2, \dots, (l-3)/2$) are not divisible by l^3 , then $x^l + y^l = z^l$ has no solution (H. S. Vandiver, *Trans. Amer. Math. Soc.*, 31 (1929)). By computation Vandiver confirmed that $x^l + y^l = z^l$ has no solution for

$l < 619$. At present, this procedure has been continued for $l \leq 125,080$ using computers by the method of D. H. Lehmer, E. Lehmer, and H. S. Vandiver, *Proc. Nat. Acad. Sci. US*, 40 (1954) (S. S. Wagstaff, *Math. Comp.*, 32 (1978)).

When the condition $(xyz, l) = 1$ or $(xyz, l) = l$ is added, we speak of Case I or Case II, respectively. The following theorems hold for Case I.

(3) If $x^l + y^l = z^l$ has a solution in Case I, then

$$B_{2m} f_{l-2m}(t) \equiv 0 \pmod{l}, \quad m = 1, 2, \dots, (l-3)/2, \tag{2}$$

holds for $-t = x/y, y/x, y/z, z/y, x/z$, and z/x , where $f_m(t) = \sum_{r=0}^{l-1} r^{m-1} t^r$, and B_m is the m th Bernoulli number. This is called **Kummer's criterion** (D. Mirimanov, 1905).

A simplification of the above result is

(4a) If $x^l + y^l = z^l$ has a solution in Case I, then

$$(2^{l-1} - 1)/l \equiv 0 \pmod{l}$$

(A. Wieferich, *J. Reine Angew. Math.*, 136 (1909)). This result created a sensation at the time of its publication. It was first shown that 1093 and 3511 are the only primes with $l < 3700$ for which the above congruence holds; it is presently known that no other l with $l \leq 6 \times 10^9$ satisfies this congruence. The criterion (4a) was gradually improved by Mirimanov (1910, 1911), P. Furtwängler (1912), Vandiver (1914), G. Frobenius (1914), F. Pollaczek (1917), T. Morishima (1931), and J. B. Rosser (1940, 1941). For example:

(4b) If $x^l + y^l = z^l$ has a solution in Case I, then

$$(m^{l-1} - 1)/l \equiv 0 \pmod{l} \tag{3}$$

holds for all m with $2 \leq m \leq 43$. By means of this result, Rosser (1941) showed for $l < 41,000,000$, and D. H. Lehmer and E. Lehmer (*Bull. Amer. Math. Soc.*, 47 (1941)) showed for $l < 253,747,889$ that $x^l + y^l = z^l$ has no solution in Case I.

We have hitherto been concerned with rational integral solutions of $x^l + y^l = z^l$. We may also consider the problem of proving or disproving that $\alpha^l + \beta^l = \gamma^l$ has no solution α, β, γ with $\alpha\beta\gamma \neq 0$ in the ring of l algebraic integers of $\mathbf{Q}(\zeta)$. Case I means the impossibility of

$$\alpha^l + \beta^l + \gamma^l = 0, \quad (\alpha\beta\gamma, l) = 1, \tag{4}$$

and Case II means the impossibility of

$$\alpha^l + \beta^l = \varepsilon \lambda^{nl} \gamma^l, \quad (\alpha\beta\gamma, l) = 1, \tag{5}$$

where n is a natural number, ε is a l unit in $\mathbf{Q}(\zeta)$, and $\lambda = 1 - \zeta$. We have the following results:

(1*) If $(h, l) = 1$, then neither equation (4) nor equation (5) has a solution (Kummer, 1850).

(2*) Under the same conditions as in statement (2), equation (4) has no solution. If we additionally restrict α, β, γ to relatively prime integers of $\mathbf{Q}(\zeta + \zeta^{-1})$ and replace λ by $(1 - \zeta)(1 - \zeta^{-1})$, then equation (5) also has no solution (Vandiver, 1929).

(4b*) If equation (4) has solutions α, β, γ in $\mathbf{Q}(\zeta)$, then congruence relation (3) holds for all m with $2 \leq m \leq 43$ (Morishima, 1934).

When l is sufficiently large, we have the results of M. Krasner (*C. R. Acad. Sci. Paris* (1934)) and Morishima (*Proc. Japan Acad.*, 11 (1935)).

Bibliographies are given in Vandiver and Wahlin [1] and Vandiver [2].

References

- [1] H. S. Vandiver and G. E. Wahlin, Algebraic numbers II, *Bull. Nat. Res. Council*, no. 62, 1928.
- [2] H. S. Vandiver, Fermat's last theorem, *Amer. Math. Monthly*, 53 (1946), 555-578.
- [3] P. Ribenboim, 13 lectures on Fermat's last theorem, Springer, 1979.

146 (XX.30) Feynman Integrals

A. Introduction

As the S -matrix or 'Green's function in quantum field theory is usually prohibitively difficult to calculate, perturbative expansions in terms of coupling constants have been employed since the beginning of the theory (\rightarrow 386 S -Matrices). R. P. Feynman (*Phys. Rev.*, 76 (1949)) invented a way of calculating the series in terms of **Feynman integrals**. His method drastically simplified the preceding method due to S. Tomonaga and J. S. Schwinger, even though, as was later shown, the two methods are theoretically equivalent (F. J. Dyson, *Phys. Rev.*, 75 (1949)). A Feynman integral is an integral associated with a **Feynman graph** according to the **Feynman rule** explained in Section B. Feynman integrals inherit the troublesome problem of divergence, and some recipe which systematically provides them with a definite meaning is needed. Such a recipe is given by the renormalization theory of Tomonaga, Schwinger, Feynman, and Dyson. A mathematically rigorous renormalization theory was given by N. N. Bogolyubov and O. S. Parasiuk (*Acta Math.*, 97 (1957)), later supplemented by K. Hepp (*Comm. Math.*

Phys., 2 (1966)). See also W. Zimmermann (*Comm. Math. Phys.*, 15 (1969)) and S. A. Anikin et al. (*Theor. Math. Phys.*, 17 (1973)). Furthermore, E. R. Speer [1] gave a mathematically convenient recipe of renormalization (under the condition that massless particles are irrelevant). Although the series expansion in coupling constants is a divergent series even after renormalization (\rightarrow 386 S-Matrices), the study of Feynman integrals has given much insight into the qualitative aspects of the S-matrix, and in particular, into its analytic structure (e.g., R. J. Eden et al. [2]). In this respect the discovery of the Landau-Nakanishi equations, which describe the location of singularities of Feynman integrals, was crucially important (L. D. Landau, N. Nakanishi, and J. Bjorken, 1959; \rightarrow Section C). Later, R. E. Cutkosky found a formula which gives important information concerning the ramification of Feynman integrals near their singularities (\rightarrow Section C). It gave impetus to J. Leray's mathematical study of Feynman integrals from the viewpoint of integration of multi-valued analytic functions (Leray, *Bull. Soc. Math. France*, 87 (1959)). Such studies were subsequently carried out by D. Fotiadi, M. Froissart, J. Lascoux, F. Pham, etc.; \rightarrow [3–5] and references cited there for this topic. An extensive study by G. Ponzano, T. Regge, Speer, and J. M. Westwater on the monodromy structure of Feynman integrals is closely related to the studies by Pham and others (\rightarrow Regge in [6] and references cited there). On the other hand, the progress of microlocal analysis has thrown new light on the Feynman integrals and has given a unified foundation to these various other studies (\rightarrow Section C; also Pham, M. Kashiwara, and T. Kawai in [6], M. Sato et al. in [6] and references cited there).

B. Definitions

First, the notion of **Feynman graphs** is introduced. A Feynman graph is sometimes called a **Feynman diagram**. A Feynman graph G consists of finitely many points (called vertices) $\{V_j\}_{j=1,\dots,n}$, finitely many 1-dimensional segments (called internal lines) $\{L_l\}_{l=1,\dots,N}$ and finitely many half-lines (called external lines) $\{L_r^e\}_{r=1,\dots,n}$, all of which are located in a 4-dimensional affine space. Each of the end-points W_l^+ and W_l^- of L_l and the endpoint of L_r^e coincide with some vertex V_j . A four-vector $p_r = (p_{r,0}, p_{r,1}, p_{r,2}, p_{r,3})$ is associated with each external line L_r^e and a constant $m_l \geq 0$ is associated with each internal line L_l . For simplicity, we usually suppose, in addition, that each internal line and each external line are

oriented, that $W_l^+ \neq W_l^-$ and that G is connected. The orientation of a line is indicated by the symbol \rightarrow . Given an orientation, we define the incidence number $[j:l]$ to be +1 or -1 according to whether L_l ends or starts from V_j . In other cases, $[j:l]$ is defined to be zero. The incidence number $[j:r]$ is defined in the same way but with L_r^e replacing L_l (Fig. 1).

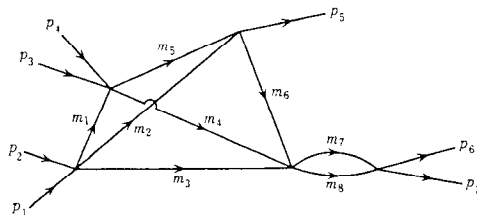


Fig. 1

In this example of a Feynman graph, the internal lines L_2 and L_4 do not intersect and this diagram should be drawn in \mathbf{R}^4 , not in \mathbf{R}^2 ; this is indicated by \nearrow . For convenience, multiple lines such as L_7 and L_8 are usually drawn in a curvilinear manner, as shown.

The **Feynman rule** associates the following integral $F_G(p)$ with each Feynman graph G :

$$F_G(p) = \int \frac{\prod_{j=1}^n \delta^4(\sum_{r=1}^n [j:r] p_r + \sum_{l=1}^N [j:l] k_l)}{\prod_{l=1}^N (k_l^2 - m_l^2 + \sqrt{-1} \ 0)} \times \sum_{i=1}^N d^4 k_i, \quad (1)$$

where $k_l^2 = k_{l,0}^2 - \sum_{v=1}^3 k_{l,v}^2$. Here $1/(k_l^2 - m_l^2 + \sqrt{-1} \ 0)$ means $\lim_{\epsilon \downarrow 0} (1/(k_l^2 - m_l^2 + \sqrt{-1} \ \epsilon))$ (\rightarrow 125 Distributions and Hyperfunctions).

Here we consider the case where the interaction Lagrangian density does not contain differential operators (i.e., direct coupling) and all relevant particles are spinless. In general, we should multiply the integrand of $F_G(p)$ by a matrix of polynomials of the p_r and k_l .

$F_G(p)$ has the form $\delta(\sum_{j,r} [j:r] p_r) f_G(p)$; we often investigate $f_G(p)$ instead of $F_G(p)$. The function $f_G(p)$ is studied on $M = \text{def} \{p \in \mathbf{R}^{4n} \mid \sum_{j,r} [j:r] p_r = 0\}$, and is called a **Feynman amplitude**. The integral (1) is not well defined as it stands because of the following problems: (a) Is the product appearing in the integrand well defined? (b) Is the integral convergent? The first problem is not serious if $m_l \neq 0$ (\rightarrow 274 Microlocal Analysis E). However, the second problem, called the **ultraviolet divergence**, is serious. The renormalization procedure is intended to overcome this difficulty. When some m_l is equal to zero, even the first problem, called the **infrared divergence**, is serious. See D. R. Yennie et al. (*Ann. Phys.*, 13 (1961)) and T. Kinoshita (*J. Math. Phys.*, 3 (1962)) for analyses of the infrared divergence. In this article we always assume for simplicity that every m_l is strictly positive, even though

such an assumption is too restrictive from the physical viewpoint.

Calculations of Feynman integrals are often done by means of the **parametric representation** of the integral (1), of the form

$$\delta\left(\sum_{j,r}[j:r]p_r\right)\int_0^1\dots\times\int_0^1\frac{\delta(1-\sum_{i=1}^N\alpha_i)\prod_{i=1}^N d\alpha_i}{U(\alpha)^2(V(p,\alpha)+\sqrt{-1}\,0)^{-N+2n'-2}},$$

where $U(\alpha)$ and $V(p,\alpha)$ are determined by the topological structure of the graph G . See [7] for the derivation of this formula and the treatment of the integral written in this form. It is useful not only for the study of the singularity structure of $f_G(p)$ but also for the study of spectral representations, etc. (\rightarrow , e.g., [7, 8]). Note that several different notations are used in the literature for the parametric representation of the integral. Hence one should be careful in referring to papers which use parametric representations.

C. Analytic Properties of Feynman Integrals

The celebrated result of Landau (*Nuclear Phys.*, 13 (1959)), Nakanishi (*Prog. Theor. Phys.*, 22 (1959)) and Bjorken (thesis, Stanford Univ., 1959) asserts that the singularities of a Feynman amplitude $f_G(p)$ are confined to the subset $L^+(G)$ (called a **positive- α Landau-Nakanishi variety**) of $M = \{p \in \mathbb{R}^{4n} \mid \sum_{j,r}[j:r]p_r = 0\}$, defined by the following set of equations (called **Landau-Nakanishi equations**), where u_r , w_j , and k_l are real four-vectors and α_i is a real number, all of which are to be eliminated to define relations among the p_r (note, however, that a positive- α Landau-Nakanishi variety is not strictly a subvariety of M , because of the constraint (2e):

$$u_r = \sum_{j=1}^{n'} [j:r]w_j \quad (r=1, \dots, n), \quad (2a)$$

$$\sum_{r=1}^n [j:r]p_r + \sum_{l=1}^N [j:l]k_l = 0 \quad (j=1, \dots, n'), \quad (2b)$$

$$\sum_{j=1}^{n'} [j:l]w_j = \alpha_l k_l \quad (l=1, \dots, N), \quad (2c)$$

$$\alpha_l(k_l^2 - m_l^2) = 0 \quad (l=1, \dots, N), \quad (2d)$$

$$\alpha_l \geq 0, \quad (2e)$$

with some

$$\alpha_l \neq 0. \quad (2f)$$

Usually a Landau-Nakanishi variety (resp., equation) is called a **Landau variety** (resp., equation) for short. The equation (2a) is conventional; (2b) represents the energy-

momentum conservation law at the vertex V_j ; (2d) corresponds to the mass-shell constraint (if $\alpha_l \geq 0$) on the internal line L_l . Since (2c) entails $\sum_l \varepsilon_l(C)\alpha_l k_l = 0$ for a closed circuit (= loop) C of G for some set of values $\varepsilon_l(C) = \pm 1$ or 0 with $\varepsilon_l(C)$ being 0 if L_l does not belong to C , the equation (2c) is usually called the closed-circuit condition. By attaching w_j to the vertex V_j and associating $\alpha_l k_l$ to the internal line L_l , we get a diagram representing a multiple collision of classical point particles, since the relations (2a)–(2f) are just the classical conditions for such a collision (S. Coleman and R. E. Norton, *Nuovo Cimento*, 38 (1965)). The fact that the Landau-Nakanishi equations admit such an interpretation is neither accidental nor superficial in view of the †macroscopic causality of the S -matrix. Note also that there is another interpretation of the Landau-Nakanishi equations, which emphasizes their resemblance to †Kirchhoff's law (Nakanishi, *Prog. Theor. Phys.*, 23 (1960)). Such a resemblance can be used to study the structure of Feynman integrals from the viewpoint of †graph theory. See [7] and the references cited there for this topic.

An important observation by Pham and Sato (1973), which opened a way to the microlocal analysis of Feynman integrals and the S -matrix, is the following: If we consider the Landau-Nakanishi equations to define a subvariety of S^*M , the †spherical cotangent bundle of M , by eliminating only w , k , and α , then the resulting variety describes the †singularity spectrum of $f_G(p)$. More precisely, $S.S.f_G(p)$ is confined to the set $(p; \sqrt{-1}u)$, where $(p; u)$ satisfies the Landau-Nakanishi equations. The rigorous proof of this statement was given by Sato et al. in [6]. The subset of S^*M or $\sqrt{-1}S^*M$ thus obtained is denoted by $\mathcal{L}^+(G)$ and is also called a positive- α Landau-Nakanishi variety. It is noteworthy that the microlocalization of the classical result of Landau, Nakanishi, and Bjorken had essentially been achieved in a less sophisticated manner by D. Iagolnitzer and H. P. Stapp (*Comm. Math. Phys.*, 14 (1969)) in the framework of S -matrix theory. The variety defined by (2a)–(2d) and (2f) is denoted by $L(G)$ or $\mathcal{L}(G)$ and is a Landau-Nakanishi variety. In a neighborhood of p^0 in $L^+(G)$, $f_G(p)$ is the boundary value of a holomorphic function $\tilde{f}_G(p)$ whose domain of definition is determined by u -vectors (\rightarrow 274 Microlocal Analysis E). Furthermore, in simple cases one can verify that $\tilde{f}_G(p)$ can be analytically continued to define a holomorphic function on the universal covering space $\widetilde{U-L(G)^C}$ of $U-L(G)^C$ for a complex neighborhood U of p^0 , where $L(G)^C$ denotes a complexification of $L(G)$. Hence we can discuss the difference of

$\tilde{f}_G(\gamma p)$ and $\tilde{f}_G(p)$, where γ denotes a loop encircling $L(G)^c$. Cutkosky (*J. Math. Phys.*, 1 (1960)) observed that it can be expressed (in simple cases) by the integral obtained by replacing $1/(k_i^2 - m_i^2 + \sqrt{-10})$ in the right-hand side of (1) by $-2\pi\sqrt{-1}\delta^+(k_i^2 - m_i^2) =_{\text{def}} -2\pi\sqrt{-1}Y(k_{i,0})\delta(k_i^2 - m_i^2)$, where $Y(k_{i,0})$ denotes the \dagger Heaviside function. A result of this type is called a **discontinuity formula** and is now obtained for the S -matrix itself in a suitably modified manner, i.e., the discontinuity formula holds beyond the framework of perturbation theory (\rightarrow 386 S -Matrices). As is mentioned in 386 S -Matrices, the discontinuity formula is closely related to Sato's conjecture on the \dagger holonomic character of the S -matrix (T. Kawai and Stapp in [6]). Actually, M. Kashiwara and Kawai (in [6]) proved that $f_G(p)$ satisfies a \dagger holonomic system of linear differential equations whose characteristic variety is confined to the extended Landau-Nakanishi variety $L(G)$. They further proved that the system has regular singularities (\rightarrow 274 Microlocal Analysis G). Their result gives, on one hand, a precise version of Regge's statement to the effect that $f_G(p)$ is a generalization of a \dagger hypergeometric function (in *Battelle Rencontres*, C. DeWitt and J. A. Wheeler (eds.), Benjamin, 1968), and, on the other hand, a rigorous proof of the fact that $\tilde{f}_G(p)$ is a Nils-son class function. This fact is closely related to the works of D. Fotiadis, J. Lascoux, Pham, and others. Kashiwara and Kawai (*Comm. Math. Phys.*, 54 (1977)) also showed that the holonomic character of the Feynman amplitude is an important clue for understanding the so-called hierarchical principle, which had been proposed and studied in connection with the \dagger Mandelstam representation by the Cambridge group (Eden et al. [2]). Furthermore, Kashiwara et al. (*Comm. Math. Phys.*, 60 (1978)) gave a useful expression of $f_G(p)$ at several physically important points by analyzing the microlocal structure of the holonomic system that $f_G(p)$ satisfies. Thus the use of \dagger (micro-) differential equations in analyzing Feynman amplitudes has turned out to be effective in understanding their singularity structures in a unified manner.

References

- [1] E. R. Speer, Generalized Feynman amplitudes, Princeton Univ. Press, 1969.
- [2] R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, The analytic S -matrix, Cambridge Univ. Press, 1966.
- [3] R. C. Hwa and V. L. Teplitz, Homology and Feynman integrals, Benjamin, 1966.

- [4] F. Pham, Introduction a l'étude topologique de singularités de Landau, Gauthier-Villars, 1967.
- [5] S. Lefschetz, Applications of algebraic topology, Springer, 1975.
- [6] Publ. Res. Inst. Math. Sci., 12, suppl. (1977) (Proc. Oji Seminar on Algebraic Analysis, 1976).
- [7] N. Nakanishi, Graph theory and Feynman integrals, Gordon & Breach, 1971. (Original in Japanese.)
- [8] I. T. Todorov, Analytic properties of Feynman diagrams in quantum field theory, Pergamon, 1971. (Original in Russian, 1966.)

147 (IX.13) Fiber Bundles

A. General Remarks

E. Stiefel [2] introduced certain \dagger diffeomorphism invariants of \dagger differentiable manifolds by considering a field of a finite number of linearly independent vectors attached to each point of a manifold; and H. Whitney [3] obtained the notion of fiber bundles as a compound idea of a manifold and such a field of tangent vectors. S. S. Chern [4] emphasized the global point of view in differential geometry by recognizing the relation between the notion of \dagger connections (due to E. Cartan) and the theory of fiber bundles. The theory of fiber bundles is also applied to various fields of mathematics, for example, the theory of \dagger Lie groups, \dagger homogeneous spaces, \dagger covering spaces, and general vector bundles, vector bundles of class C^r , or analytic vector bundles.

Homological properties of fiber bundles are studied by means of \dagger spectral sequences, and cohomology structures of several homogeneous spaces and several characteristic classes are determined explicitly by means of \dagger cohomology operations. Also, the group $K(X)$, formed by equivalence classes of vector bundles over a finite \dagger CW complex X , is a \dagger generalized cohomology group, treated in $\dagger K$ -theory, in which further development is expected (\rightarrow 237 K -Theory).

B. Definitions

Let E, B, F be topological spaces, $p: E \rightarrow B$ a continuous mapping, and G an \dagger effective left topological \dagger transformation group of F . If there exist an \dagger open covering $\{U_\alpha\}$ ($\alpha \in \Lambda$) of B and a homeomorphism $\varphi_\alpha: U_\alpha \times F \approx p^{-1}(U_\alpha)$ for each $\alpha \in \Lambda$ having the following three properties, then the system $(E, p, B, F, G, U_\alpha, \varphi_\alpha)$ is

called a **coordinate bundle**: (1) $p\varphi_\alpha(b, y) = b$ ($b \in U_\alpha, y \in F$). (2) Define $\varphi_{\alpha, b}: F \approx p^{-1}(b)$ ($b \in U_\alpha$) by $\varphi_{\alpha, b}(y) = \varphi_\alpha(b, y)$; then $g_{\beta\alpha}(b) = \varphi_{\beta, b}^{-1} \circ \varphi_{\alpha, b} \in G$ for $b \in U_\alpha \cap U_\beta$. (3) $g_{\beta\alpha}: U_\alpha \cap U_\beta \rightarrow G$ is continuous. We say that this bundle is **equivalent** to a coordinate bundle $(E, p, B, F, G, U'_\mu, \varphi'_\mu)$ if $\bar{g}_{\mu\alpha}(b) = \varphi'^{-1}_{\mu, b} \circ \varphi_{\alpha, b} \in G$ ($b \in U_\alpha \cap U'_\mu$) and $\bar{g}_{\mu\alpha}: U_\alpha \cap U'_\mu \rightarrow G$ is continuous. An equivalence class $\xi = (E, p, B, F, G)$ of coordinate bundles is called a **fiber bundle** (or **G-bundle**), and E is called the **total space** (or **bundle space**), p the **projection**, B the **base space**, F the **fiber**, and G the **bundle group** (or **structure group**). Also, U_α of a coordinate bundle $(E, p, B, F, G, U_\alpha, \varphi_\alpha)$ belonging to the class ξ is called the **coordinate neighborhood**, φ_α the **coordinate function**, and $g_{\beta\alpha}$ the **coordinate transformation** (or **transition function**).

Let $\xi = (E, p, B, F, G)$ and $\xi' = (E', p', B', F, G)$ be two fiber bundles with the same fiber and group. A continuous mapping $\Psi: E \rightarrow E'$ is called a **bundle mapping** (**bundle map**) from ξ to ξ' if the following two conditions are satisfied: (1) There is a continuous mapping $\psi: B \rightarrow B'$ with $p' \circ \Psi = \Psi \circ p$. (2) $\psi_{\mu\alpha}(b) = \varphi'^{-1}_{\mu, b'} \circ \Psi \circ \varphi_{\alpha, b} \in G$ ($b \in U_\alpha \cap \psi^{-1}(V'_\mu)$, $b' = \psi(b)$), and $\psi_{\mu\alpha}: U_\alpha \cap \psi^{-1}(V'_\mu) \rightarrow G$ is continuous, where $\{U_\alpha, \varphi_\alpha\}$ and $\{V'_\mu, \varphi'_\mu\}$ are pairs of coordinate neighborhoods and functions of ξ and ξ' , respectively. Moreover, if ψ is a homeomorphism, then Ψ is also a homeomorphism and Ψ^{-1} is a bundle mapping.

Let $\xi = (E, p, B, F, G)$ and $\xi' = (E', p', B', F, G)$ be two fiber bundles with the same base space, fiber, and group. If there is a bundle mapping $\Psi: E \rightarrow E'$ such that $\psi: B \rightarrow B$ as described before is the identity mapping, then we say that ξ is **equivalent** (or **isomorphic**) to ξ' and write $\xi \equiv \xi'$. Take the same coordinate neighborhoods $\{U_\alpha\}$, and let $g_{\beta\alpha}$ and $g'_{\beta\alpha}$ be the coordinate transformations of ξ and ξ' , respectively. Then $\xi \equiv \xi'$ if and only if there are continuous mappings $\lambda_\alpha: U_\alpha \rightarrow G$ with $g'_{\beta\alpha}(b) = \lambda_\beta(b)g_{\beta\alpha}(b)\lambda_\alpha(b)^{-1}$ ($b \in U_\alpha \cap U_\beta$).

For a system $\{g_{\beta\alpha}\}$ of coordinate transformations of a fiber bundle, we have $g_{\gamma\beta}(b)g_{\beta\alpha}(b) = g_{\gamma\alpha}(b)$ ($b \in U_\alpha \cap U_\beta \cap U_\gamma$). Conversely, given an open covering $\{U_\alpha\}$ and a system of $g_{\beta\alpha}: U_\alpha \cap U_\beta \rightarrow G$ satisfying this condition, there is a unique G -bundle (E, p, B, F, G) with $\{g_{\beta\alpha}\}$ as a system of coordinate transformations. Actually, E is the \dagger identification space of $\tilde{E} = \{(b, y, \alpha) | b \in U_\alpha\} \subset B \times F \times \Lambda$ obtained by identifying two points $(b, y, \alpha), (b', y', \beta)$ with $b = b', y' = g_{\beta\alpha}(b) \cdot y$, and p is defined by $p\{(b, y, \alpha)\} = b$, where the index set $\Lambda = \{\alpha\}$ is considered a discrete space.

C. Principal Fiber Bundles

A fiber bundle $\eta = (P, q, B, G, G)$ is called a **principal fiber bundle** (or simply **principal**

bundle) if G operates on G by left translations. This is also defined by the following conditions: G is a right \dagger topological transformation group of P , and there exist an open covering $\{U\}$ of B and homeomorphisms $\varphi: U \times G \approx q^{-1}(U)$ with $q\varphi(b, g) = b$, $\varphi(b, g) \cdot g' = \varphi(b, gg')$ ($b \in U; g, g' \in G$). A bundle mapping $\Psi: P \rightarrow P'$ between two principal bundles $\eta = (P, q, B, G)$ and $\eta' = (P', q', B', G)$ is also defined as a continuous mapping Ψ with $\Psi(x \cdot g) = \Psi(x) \cdot g$.

D. Associated Fiber Bundles

Let $\eta = (P, q, B, G)$ be a principal bundle, and let F be a topological space having G as an effective left topological transformation group. Then G is a right topological transformation group of the product space $P \times F$ by $(x, y) \cdot g = (x \cdot g, g^{-1} \cdot y)$ ($x \in P, y \in F, g \in G$). Consider the \dagger orbit space $P \times_G F = (P \times F)/G$, and define the continuous mapping $p: P \times_G F \rightarrow B$ by $p\{(x, y)\} = q(x)$. Then $\eta \times_G F = (P \times_G F, p, B, F, G)$ is a fiber bundle, called the **associated fiber bundle** of η with fiber F . On the other hand, η is called an **associated principal bundle** of $\xi = (E, p, B, F, G)$ if $\xi \equiv \eta \times_G F$. A principal bundle η having the same coordinate transformations as ξ is an associated principal bundle of ξ , and two fiber bundles are equivalent if and only if their associated principal bundles are equivalent. Therefore, given a fiber bundle ξ , there exists a principal bundle η such that $\xi = \eta \times_G F$.

E. Examples of Fiber Bundles

(1) **Product bundle**. $(B \times F, p_1, B, F, G)$, where p_1 , the projection of the product space, is called a **product bundle** if there is just one coordinate neighborhood B and the coordinate function is the identity mapping of $B \times F$. A bundle that is equivalent to a product bundle is called a **trivial bundle**.

(2) A \dagger covering (\tilde{Y}, p, Y) is a fiber bundle whose fiber is the discrete space $p^{-1}(y_0)$ ($y_0 \in Y$), and the structure group is a factor group of the \dagger fundamental group $\pi_1(Y, y_0)$. In particular, a \dagger regular covering is a principal bundle.

(3) **Hopf bundle**. Let Λ be the real number field \mathbf{R} , the complex number field \mathbf{C} , or the quaternion field \mathbf{H} , $\lambda = \dim_{\mathbf{R}} \Lambda$, and Λ^{n+1} the $(n+1)$ -dimensional linear space over Λ . Identify two points $(z_0, \dots, z_n), (z'_0, \dots, z'_n)$ of the subspace $\Lambda^{n+1} - \{0\}$ (0 is the origin) if there is a $z \in \Lambda$ such that $z_i = z'_i z$ ($i = 0, \dots, n$). Then we obtain the identification space $P^n(\Lambda)$, called the n -dimensional **projective space over Λ** . Let $S^n_\Lambda (= S^{\lambda(n+1)-1})$, the $(\lambda(n+1)-1)$ -sphere be the unit sphere in Λ^{n+1} . Then S^n_Λ is the topological

transformation group of S_Λ^n by the product of Λ , and the orbit space S_Λ^n/S_Λ^0 is $P^n(\Lambda)$. Furthermore, $(S_\Lambda^n, q, P^n(\Lambda), S_\Lambda^0)$ (q is the projection) is a principal bundle called the **Hopf bundle** (or **Hopf fibering**). These comments are valid also for $n = \infty$. When $n = 1$, $P^1(\Lambda)$ is homeomorphic to S^λ , and the Hopf bundle is $(S^{2\lambda-1}, q, S^\lambda, S^{\lambda-1})$ ($\lambda = 1, 2, 4$). A Hopf bundle is defined similarly for $\lambda = 8$ using the \dagger Cayley algebra, and the projection $q: S^{2\lambda-1} \rightarrow S^\lambda$ ($\lambda = 2, 4, 8$) is the **Hopf mapping** (**Hopf map**).

(4) Let G be a topological group, H its closed subgroup, and $r: G \rightarrow G/H$, $r(g) = gH$ the natural projection. If there exist a neighborhood U of $r(H) \in G/H$ and a continuous mapping $f: U \rightarrow G$ such that $r \circ f$ is the identity mapping, then we say that H has a **local cross section** f in G , and $(G/K, p, G/H, H/K, H/K_0)$ is a fiber bundle for any closed subgroup K of H (where p is the natural projection $gK \rightarrow gH$ and K_0 is the largest normal subgroup of H contained in K). The associated principal bundle of the latter fiber bundle is $(G/K_0, p, G/H, H/K_0)$. Any closed subgroup H of a \dagger Lie group G has a local cross section in G ; hence the above bundles can be obtained.

F. Vector Bundles

A system $\xi = (E, p, B)$ of topological spaces E, B and a continuous mapping $p: E \rightarrow B$ is called an n -dimensional real **vector bundle** if the following two conditions are satisfied: (1) $p^{-1}(b)$ is a real vector space for each $b \in B$. (2) There exist an open covering $\{U_\alpha\}$ ($\alpha \in \Lambda$) of B and a coordinate function $\varphi_\alpha: U_\alpha \times F \approx P^{-1}(U_\alpha)$ for each $\alpha \in \Lambda$, where $F = \mathbf{R}^n$; furthermore, the $\varphi_{\alpha,b}: \mathbf{R}^n \approx p^{-1}(b)$ are isomorphisms of vector spaces. In this case, $g_{\beta\alpha}(b) = \varphi_{\beta,b}^{-1} \circ \varphi_{\alpha,b}: \mathbf{R}^n \approx \mathbf{R}^n$ ($b \in U_\alpha \cap U_\beta$) is an element of the \dagger general linear group $GL(n, \mathbf{R})$. Hence a vector bundle is a fiber bundle with fiber \mathbf{R}^n and group $GL(n, \mathbf{R})$, and the converse is also true. A 1-dimensional vector bundle is called a **line bundle**. A vector bundle $\xi' = (E', p', B)$ is called a **subbundle** of a vector bundle $\xi = (E, p, B)$ if $E' \subset E$, $p|_{E'} = p'$, and $p^{-1}(b)$ is a vector subspace of $p^{-1}(b)$ for each $b \in B$.

Let ξ_1 and ξ_2 be two vector bundles of dimension n_1 and n_2 with the same base space B . Let E be the union of the direct sum $p_1^{-1}(b) + p_2^{-1}(b)$ for $b \in B$, and define $p: E \rightarrow B$ by $p(p_1^{-1}(b) + p_2^{-1}(b)) = b$. Take the same coordinate neighborhoods U_α for ξ_1 and ξ_2 , and define $\varphi_\alpha: U_\alpha \times \mathbf{R}^{n_1+n_2} \rightarrow p^{-1}(U_\alpha)$ by $\varphi_\alpha(b, y) = (\varphi_{\alpha,b}^1 + \varphi_{\alpha,b}^2)(y)$ ($y \in \mathbf{R}^{n_1+n_2} = \mathbf{R}^{n_1} + \mathbf{R}^{n_2}$), where the $\varphi_\alpha^i: U_\alpha \times \mathbf{R}^{n_i} \approx p_i^{-1}(U_\alpha)$ are the coordinate functions of ξ_i . Then E is topologized by taking the family $\{\varphi_\alpha(O)\}$ (O is open in $U_\alpha \times \mathbf{R}^{n_1+n_2}$) as the \dagger open base, and we obtain

an $(n_1 + n_2)$ -dimensional real vector bundle (E, p, B) , denoted by $\xi_1 \oplus \xi_2$ and called the **Whitney sum** of ξ_1 and ξ_2 . Similarly, we can define the **tensor product** $\xi_1 \otimes \xi_2$, the **p -fold exterior power** $\wedge^p \xi$ (or **bundle** $\xi^{(p)}$ of p -vectors), and the **bundle of homomorphisms** $\text{Hom}(\xi_1, \xi_2)$, of dimension $n_1 n_2$, $\binom{n}{p}$, and $n_1 n_2$, respectively, using the tensor product $\mathbf{R}^{n_1} \otimes \mathbf{R}^{n_2} = \mathbf{R}^{n_1 n_2}$, the p -fold \dagger exterior power $\wedge^p \mathbf{R}^n = \mathbf{R}^{\binom{n}{p}}$ (the space $(\mathbf{R}^n)^{(p)}$ of p -vectors in \mathbf{R}^n), and the space of homomorphisms $\text{Hom}(\mathbf{R}^{n_1}, \mathbf{R}^{n_2}) = \mathbf{R}^{n_1 n_2}$. (For the last one, we use $\text{Hom}((\varphi_{\alpha,b}^1)^{-1}, \varphi_{\alpha,b}^2)$ instead of $\text{Hom}(\varphi_{\alpha,b}^1, \varphi_{\alpha,b}^2)$.) $\text{Hom}(\xi, \varepsilon^1) = \xi^*$ is called the **dual (vector) bundle** of ξ , where ε^1 is the trivial line bundle. If we use coordinate transformations, \oplus , \otimes , \wedge^p , and ξ^* are obtained by the direct sum, \dagger Kronecker product, matrix of $\dagger p$ -minors, and \dagger transpose of matrices, respectively. If ξ_2 is a subbundle of ξ_1 , the **quotient bundle** ξ_1/ξ_2 of dimension $n_1 - n_2$ is defined by using the quotient vector space $\mathbf{R}^{n_1}/\mathbf{R}^{n_2} = \mathbf{R}^{n_1 - n_2}$, and $\xi_2 \oplus (\xi_1/\xi_2)$ is equivalent to ξ_1 . These operations preserve the equivalence relation of bundles. Also, \oplus and \otimes are commutative up to equivalence and satisfy the associative and distributive laws. For each ξ having a finite-dimensional \dagger CW complex as base space, there is a ξ' such that $\xi \oplus \xi'$ is trivial.

Using the complex number field \mathbf{C} or the quaternion field \mathbf{H} instead of the real number field \mathbf{R} , we can define similarly the **complex vector bundle** or the **quaternion vector bundle** and the operations \oplus , \otimes , etc. For a complex vector bundle ξ , the **complex conjugate bundle** $\bar{\xi}$ is defined by the complex conjugate of matrices.

(5) Tangent bundles, tensor bundles. Let M be an n -dimensional \dagger differentiable manifold of class C^r . Consider the \dagger tangent vector space $T_p(M)$ at $p \in M$, set $T(M) = \bigcup_{p \in M} T_p(M)$, and define $\pi: T(M) \rightarrow M$ by $\pi(T_p(M)) = p$. For a \dagger coordinate neighborhood U_p of p with local coordinate system (x_1, \dots, x_n) , each point of $\pi^{-1}(U_p)$ is represented by $\sum_{i=1}^n f_i \partial/\partial x_i$, and $\pi^{-1}(U_p)$ has a coordinate system $(x_1, \dots, x_n, f_1, \dots, f_n)$. Hence $T(M)$ is a C^{r-1} -manifold, and $\mathfrak{T}(M) = (T(M), \pi, M, \mathbf{R}^n, GL(n, \mathbf{R}))$ is an n -dimensional real vector bundle. $\mathfrak{T}(M)$ is called the **tangent (vector) bundle**, its dual bundle $\mathfrak{T}^*(M)$ the **cotangent (vector) bundle**, and the tensor product $\mathfrak{T}(M) \otimes \dots \otimes \mathfrak{T}^*(M) \otimes \dots$ a **tensor bundle** of M . The line bundle $\wedge^n \mathfrak{T}^*(M)$ is called the **canonical bundle** of M .

For a \dagger complex manifold M , $T(M)$ is a complex manifold and $\mathfrak{T}(M)$ is a complex vector bundle. Therefore these bundles are defined as complex bundles.

(6) Tangent r -frame bundle. In the preceding example, the space of all \dagger tangent r -frames of M is a bundle space with base space M and

group $GL(n, \mathbf{R})$. It is called the **tangent r -frame bundle** (or the bundle of tangent r -frames) of M .

G. The Classification Problem

For a fiber bundle $\xi = (E, p, B, F, G)$ and a continuous mapping $\psi: B' \rightarrow B$, consider the subspace $E' = \{(x, b') \in E \times B' \mid p(x) = \psi(b')\}$ of $E \times B'$ and the projections $p': E' \rightarrow B'$ and $\Psi: E' \rightarrow E$. Then $\psi^* \xi = (E', p', B', F, G)$ is a fiber bundle, and Ψ is a bundle mapping from $\psi^* \xi$ to ξ ; $\psi^* \xi$ is called the **induced bundle** of ξ by ψ . Let $\{U_\alpha\}$ and $\{g_{\beta\alpha}\}$ be the systems of coordinate neighborhoods and transformations of ξ . Then $\{\psi^{-1}(U_\alpha)\}$ and $\{g_{\beta\alpha} \circ \psi\}$ are corresponding systems of $\psi^* \xi$. If $\Psi: E' \rightarrow E$ is a bundle mapping from ξ' to ξ having $\psi: B' \rightarrow B$ as the mapping of base spaces, then $\xi' \equiv \psi^* \xi$. Also, we have $\psi^* \xi_1 \equiv \psi^* \xi_2$ if $\xi_1 \equiv \xi_2$; $(\psi \circ \psi')^* \xi \equiv \psi'^*(\psi^* \xi)$. If ξ is a principal bundle, then $\psi^* \xi$ is also principal, and $\psi^*(\eta \times_G F) \equiv (\psi^* \eta) \times_G F$. For a \dagger paracompact space B' , $\psi_1^* \xi \equiv \psi_2^* \xi$ if $\psi_1, \psi_2: B' \rightarrow B$ are \dagger homotopic.

For a topological group G , a principal bundle $\xi(n, G) = (E(n, G), p, B(n, G), G)$ is called an **n -universal bundle** if $E(n, G)$ is $\dagger n$ -connected ($n \leq \infty$); its base space $B(n, G)$ is called an **n -classifying space** of G . In particular, $\xi(\infty, G) = \xi_G = (E_G, p, B_G, G)$ is called simply a **universal bundle** and B_G a **classifying space** of G . Then we have the **classification theorem**: Let B be a CW complex with $\dim B \leq n$; then the set of equivalence classes of principal G -bundles with base space B is in one-to-one correspondence with the \dagger homotopy set $\pi(B; B(n, G))$ of continuous mappings of B into $B(n, G)$. Such a correspondence is given by associating with the induced bundle $\psi^* \xi(n, G)$ a continuous mapping $\psi: B \rightarrow B(n, G)$, called the **characteristic mapping or classifying mapping (characteristic map or classifying map)** of $\psi^* \xi(n, G)$. Furthermore, if G is an effective left topological transformation group of F , the set of equivalence classes of G -bundles with base space B and fiber F is in one-to-one correspondence with $\pi(B; B(n, G))$. The correspondence is given by associating $\psi^*(\xi(n, G) \times_G F)$ to ψ .

H. Construction of Universal Bundles

For an arbitrary topological group G , J. W. Milnor [6] constructed a universal bundle (E_G, p, B_G, G) in the following manner. The **join** $E_G = G \circ \dots \circ G \circ \dots$ of countably infinite copies of G is defined as follows: A point e of E_G is the symbol $t_1 g_{i_1} \oplus \dots \oplus t_m g_{i_m}$ ($1 \leq i_1 < i_2 < \dots < i_m, m = 1, 2, 3, \dots$), where t_1, \dots, t_m are real numbers satisfying $t_{i_k} \geq 0, t_{i_1} + \dots + t_{i_m} = 1$,

and g_{i_k} is an element of the i_k th copy of G , $k = 1, \dots, m$. Here we can omit $t_{i_k} g_{i_k}$ if $t_{i_k} = 0$. Regard E_G as a topological space with a weak topology such that the coordinate functions $e \mapsto t_i, e \mapsto g_i$ are continuous. Define the right action of G on E_G by $(t_{i_1} g_{i_1} \oplus \dots \oplus t_{i_m} g_{i_m}) \cdot g = t_{i_1} (g_{i_1} \cdot g) \oplus \dots \oplus t_{i_m} (g_{i_m} \cdot g)$. Let B_G and $p: E_G \rightarrow B_G$ be the identification space of E_G by this action of G and the identification mapping. Then (E_G, p, B_G, G) is a **universal bundle** for G , and B_G is called the **classifying space** for G . E_G and B_G are sometimes written as EG and BG . In particular, a classifying space B_G is a countable CW complex for any countable CW group G (i.e., a topological group that is a countable CW complex such that the mapping $g \mapsto g^{-1}$ of G into G and the product mapping $G \times G \rightarrow G$ are both cellular). The following examples of classifying spaces for Lie groups are also useful. Note that every CW complex B_G of a given G has the same \dagger homotopy type.

I. Examples of Universal Bundles

(1) G is either $O(n)$, $U(n)$, or $Sp(n)$: Let Λ and λ be as in (3) of Section E. According as Λ is \mathbf{R} , \mathbf{C} , or \mathbf{H} , we let $U(n, \Lambda)$ be the \dagger orthogonal group $O(n)$, the \dagger unitary group $U(n)$, or the \dagger symplectic group $Sp(n)$. Then the \dagger Stiefel manifold $V_{m+n, m}(\Lambda) = U(m+n, \Lambda)/I_m \times U(n, \Lambda)$ (I_m is the unit element of $U(m, \Lambda)$) is $(\lambda(n+1) - 2)$ -connected. Hence the principal bundle $\xi(\lambda(n+1) - 2, U(m+n, \Lambda)) = (V_{m+n, m}(\Lambda), M_{m+n, m}(\Lambda), U(m, \Lambda))$ from (4) of Section E is a $(\lambda(n+1) - 2)$ -universal bundle of $U(m, \Lambda)$, where the base space $M_{m+n, m}(\Lambda) = U(m+n, \Lambda)/U(m, \Lambda) \times U(n, \Lambda)$ is the \dagger Grassmann manifold.

(2) G is either $O(\infty)$, $U(\infty)$, or $Sp(\infty)$. The examples in (1) are valid for $m, n = \infty$. Consider the \dagger inductive limit group $U(\infty, \Lambda) = \bigcup_n U(n, \Lambda)$ under the natural inclusion $U(n, \Lambda) \subset U(n+1, \Lambda)$, and supply the **infinite classical group** $U(\infty, \Lambda)$ with the weak topology (this means that a set O of $U(\infty, \Lambda)$ is open if and only if each $O \cap U(n, \Lambda)$ is open in $U(n, \Lambda)$). Then the **infinite Stiefel manifold** $V_{m+n, m}(\Lambda)$ and the **infinite Grassmann manifold** $M_{m+n, m}(\Lambda)$ ($m = \infty$ or $n = \infty$) are defined as before, and we have

$$M_{\infty, m}(\Lambda) = \bigcup_n M_{m+n, m}(\Lambda),$$

and so on. Furthermore, these manifolds are CW complexes, and $V_{\infty, m}(\Lambda)$ ($m \leq \infty$) is ∞ -connected. Although $U(\infty, \Lambda)$ is not a Lie group, $U(m, \Lambda) \times U(n, \Lambda)$ has a local cross section in $U(m+n, \Lambda)$ for $m, n \leq \infty$ [7]. Therefore, setting $n = \infty$ in (1), $\xi(\infty, U(m, \Lambda))$ is a universal bundle of $U(m, \Lambda)$, and the infinite Grassmann manifold $M_{\infty, m}(\Lambda)$ is a classifying

space $B_{U(m,\Lambda)}$. Also, $\xi(\lambda(n+1)-2, U(\infty, \Lambda))$ in (1) is a $(\lambda(n+1)-2)$ -universal bundle of $U(\infty, \Lambda)$ ($n \leq \infty$).

(3) G is either $SO(m)$ or a general Lie group. For the rotation group $SO(m)$, we have $\xi(n-1, SO(n)) = (V_{m+n,m}(\mathbf{R}), p, \tilde{M}_{m+n,m}, SO(m))$ and $B_{SO(m)} = \tilde{M}_{\infty,m}$, where $\tilde{M}_{m+n,m} = SO(m+n)/SO(m) \times SO(n)$ is the oriented Grassmann manifold. For any compact Lie group G , we have $\xi(n-1, G) = (V_{m+n,m}(\mathbf{R}), p, O(m+n)/G \times O(n), G)$, where $G \subset O(m)$. For any connected Lie group G , we have $\xi(n, G) = \xi(n, G_1) \times_{G_1} G$, where G_1 is the maximum compact subgroup of G (since G/G_1 is homeomorphic to a Euclidean space, $\xi(n, G)$ reduces to $\xi(n, G_1)$).

J. Reduction of Fiber Bundles

Let G be a topological group and H its closed subgroup. We say that the structure group of a G -bundle ξ is **reduced** to H if ξ is equivalent to a G -bundle whose coordinate transformations take values in H . For a principal H -bundle $\eta_0 = (P, q, B, H)$, the associated H -bundle $\eta_0 \times_H G = (P \times_H G, p, B, G)$ with fiber G is defined, where H operates on G by the product of G ; it is also a principal G -bundle if we define an operation of G on $P \times_H G$ by $\{(x, g)\} \cdot g' = \{(x, gg')\}$. For a principal G -bundle η , we say that η is **reducible** to an H -bundle if there is a principal H -bundle η_0 with $\eta = \eta_0 \times_H G$, and we call η_0 a **reduced bundle** of η . It is easy to see that the group of a G -bundle ξ is reducible to H if and only if the associated principal G -bundle of ξ is reducible to H . Also, if η_0 is a reduced bundle of η , then $\psi^\# \eta_0$ is a reduced bundle of $\psi^\# \eta$.

Now, assume that H has a local cross section in G and G/H is ∞ -connected. Then for an n -universal bundle $\xi(n, H)$ of H , $\xi(n, H) \times_H G$ is an n -universal bundle of G (n -connectedness of $E(n, H) \times_H G$ is shown by the homotopy exact sequence of fiber spaces). Therefore, by the classification theorem, the group of any G -bundle is reducible to H , and the equivalence classes of G -bundles are in one-to-one correspondence with those of H -bundles.

(1) A G -bundle is trivial if and only if its group is reducible to e (identity element). A $2n$ -dimensional differentiable manifold M of class C^∞ has an almost complex structure if and only if the group of the tangent bundle $\mathfrak{T}(M)$ is reducible to $GL(n, \mathbf{C})$, i.e., $\mathfrak{T}(M)$ is considered as an n -dimensional complex vector bundle.

(2) Since $GL(n, \mathbf{R}) \approx O(n) \times \mathbf{R}^{n(n+1)/2}$ and $GL(n, \mathbf{C}) \approx U(n) \times \mathbf{R}^{n^2}$, n -dimensional real (complex) vector bundles can be considered as $O(n)$ ($U(n)$)-bundles with fiber \mathbf{R}^n (\mathbf{C}^n).

K. Homotopy and Homology Theory of Bundles

Since a fiber bundle is a locally trivial fiber space, the exact sequence and the spectral sequence of fiber spaces (\rightarrow 148 Fiber Spaces) are applicable to fiber bundles. For example, the cohomology structures of homogeneous spaces of classical groups have been determined by A. Borel, J.-P. Serre, and others.

(1) Characteristic class. For a classifying space B_G of a topological group G , we have an isomorphism $\pi_n(B_G) \cong \pi_{n-1}(G)$ of homotopy groups and the following classification theorem of fiber bundles over the n -sphere S^n . The set of the equivalence classes of principal G -bundles or G -bundles with fiber F over the base space S^n is in one-to-one correspondence with the set $\pi_{n-1}(G)/\pi_0(G)$ of equivalence classes under the operation of G on $\pi_{n-1}(G)$ given by the inner automorphisms of G ; such a correspondence is given by associating with each principal G -bundle $\eta = (P, q, S^n, G)$ the class (called the **characteristic class** of η) containing the image $\Delta(i_n)$ of a generator $i_n \in \pi_n(S^n)$ by the homomorphism $\Delta: \pi_n(S^n) \cong \pi_n(P, G) \rightarrow \pi_{n-1}(G)$. Take U_1 and U_2 (the open sets of S^n such that the last coordinates t_{n+1} are $> -1/2$ and $< 1/2$, respectively) as coordinate neighborhoods of η . Then the restriction $T = g_{12}|S^{n-1}$ represents the characteristic class of η , where $g_{12}: U_1 \cap U_2 \rightarrow G$ is the coordinate transformation and S^{n-1} is the equator of S^n .

(2) For the principal bundle $\eta = (SO(n+1), q, S^n, SO(n))$, the mapping $T: S^{n-1} \rightarrow SO(n)$ is given by

$$T(t_1, \dots, t_n) = (I_n - 2(t_i t_j)) \begin{pmatrix} I_{n-1} & 0 \\ 0 & -1 \end{pmatrix}$$

(I_n is the unit matrix of degree n). Hence the mapping degree of the composite $q' \circ T: S^{n-1} \rightarrow S^{n-1}$ (of T and the natural projection $q': SO(n) \rightarrow S^{n-1}$) is equal to 0 if n is odd and 2 if n is even. From this fact and the homotopy exact sequence, we have

$$\pi_n(V_{m+n,m}(\mathbf{R})) = \begin{cases} \mathbf{Z} & \text{if } m=1 \text{ or } n \text{ is even} \\ \mathbf{Z}_2 = \mathbf{Z}/2\mathbf{Z} & \text{if } m>1 \text{ and } n \text{ is odd} \end{cases}$$

for the real Stiefel manifold $V_{m+n,m}(\mathbf{R})$, which is $(n-1)$ -connected.

(3) Sphere bundles. An $O(n+1)$ -bundle with fiber S^n is called an **n -sphere bundle**. The set of equivalence classes of n -sphere bundles with base space S^m is in one-to-one correspondence with $\pi_{m-1}(O(n+1))/\pi_0(O(n+1))$. For example, any 1-sphere bundle over S^m ($m \geq 3$) and any n -sphere bundle over S^3 is trivial. Every 3-sphere bundle over S^4 is equivalent to one of $\{\xi_{m,n} | m \text{ an integer, } n \text{ a positive integer}\}$, where

$\xi_{m,n}$ is defined as follows: Let $\rho, \sigma: S^3 \rightarrow O(4)$ be defined by $\rho(q)q' = qq'q^{-1}$, $\sigma(q)q' = qq'$ (q, q' are \dagger quaternions of norm 1). Then these mappings represent generators of $\pi_3(O(4)) \cong \pi_3(S^3 \times S^3) \cong \mathbf{Z} + \mathbf{Z}$, and $\xi_{m,n}$ is the 3-sphere bundle over S^4 corresponding to the element $m\{\rho\} + n\{\sigma\} \in \pi_3(O(4))$ (i.e., to the mapping $f_{m,n}: S^3 \rightarrow O(4)$ defined by $f_{m,n}(q)(q') = q^{m+n}q'q^{-m}$). (Here we use the fact that the operation of the element $r \in O(4)$ ($r(q) = q^{-1}$) is given by $r\rho r^{-1} = \rho$, $r\sigma r^{-1} = \rho\sigma^{-1}$.)

L. Cross Sections

For a fiber bundle $\xi = (E, p, B, F, G)$, a cross section $f: B_0 \rightarrow E$ over a subspace $B_0 (\subset B)$ is a continuous mapping such that $p \circ f$ is the identity mapping of B_0 ; a cross section over B is called a **cross section** of ξ . A bundle ξ is trivial if and only if the associated principal bundle of ξ has a cross section. More generally, given a principal bundle $\eta = (P, q, B, G)$ and a closed subgroup H having a local cross section in G , η is reducible to H if and only if the associated bundle $\eta(\times_G(G/H) = (P/H, q', B), G/H)$ with fiber G/H has a cross section.

Suppose that the base space B of the fiber bundle $\xi = (E, p, B, F, G)$ is a \dagger polyhedron. We denote the $\dagger r$ -skeleton of B by B^r and consider the problem of extending cross sections $f_r: B^r \rightarrow E$ successively for $r = 0, 1, \dots$. Clearly, there is a cross section f_0 . For each r -simplex σ of B , we have $(p^{-1}(\sigma), p, \sigma, F) \equiv (\sigma \times F, p_1, \sigma, F)$ since σ is \dagger contractible. Hence there is a bundle mapping $\varphi_\sigma: \sigma \times F \approx p^{-1}(\sigma)$ with $p \circ \varphi_\sigma = p_1$. Assume the existence of a cross section $f_{r-1}: B^{r-1} \rightarrow E$, and consider the mapping $h_\sigma = p_2 \circ \varphi_\sigma^{-1} \circ (f_{r-1}|_\sigma): \dot{\sigma} \rightarrow F$ ($p_2: \sigma \times F \rightarrow F$ is the projection and $\dot{\sigma}$ is the boundary of σ). Then if h_σ is extensible to $h_\sigma: \sigma \rightarrow F$, an extended cross section $f_\sigma: \sigma \rightarrow E$ of $f_{r-1}|_\sigma$ is defined by $f_\sigma(b) = \varphi_\sigma(b, h_\sigma(b))$ ($b \in \sigma$), and the extension $f_r: B^r \rightarrow E$ of f_{r-1} is defined by $f_r|_\sigma = f_\sigma$, $f_r|_{B^{r-1}} = f_{r-1}$. If $\pi_{r-1}(F) = 0$, for example, there is an extension h_σ of h_σ since $(\sigma, \dot{\sigma}) \approx (V^r, S^{r-1})$, and f_{r-1} is extensible to a cross section f_r .

Now assume that the base space B of a G -bundle $\xi = (E, p, B, F, G)$ is an \dagger arcwise connected polyhedron and F is $\dagger(n-1)$ -connected. Then there is a cross section $f: B^n \rightarrow E$ constructed by the stepwise method of the previous paragraph. But if $\pi_n(F) \neq 0$, we have an obstruction to extending f over B^{n+1} . Now we explain how to measure this obstruction. Suppose that F is $\dagger n$ -simple. Then for each $(n+1)$ -simplex σ of B , the mapping $h_\sigma: \dot{\sigma} \rightarrow F$, defined by f as in the previous paragraph, determines a unique element $c(f)(\sigma)$ of the homotopy group $\pi_n(F)$. Hence we have a \dagger cochain $c(f) \in C^{n+1}(B; \pi_n(F))$, and f is extensible to a cross

section over B^{n+1} if and only if $c(f) = 0$. Thus there is a cross section over B^{n+1} if and only if the set of $c(f)$ for every cross section f over B^n contains the cochain 0; $\{c(f)\}$ is considered as a measure of the obstruction.

Let $w: I \rightarrow B$ be a \dagger path. We consider the space $I \times F$ and the canonical projection $p_1: I \times F \rightarrow I$. Then there is a bundle mapping $\Omega: I \times F \rightarrow E$ with $p \circ \Omega = w \circ p_1$, since $w^\# \xi \equiv (I \times F, p_1, I, F)$; and a homeomorphism $w_\#: F \approx F$ is defined by $w_\# = \varphi_{0, b_0}^{-1} \circ \Omega_0 \circ \Omega_1^{-1} \circ \varphi_{1, b_1}$, where $b_\varepsilon = w(\varepsilon)$ ($\varepsilon = 0, 1$), $\varphi_\varepsilon: U_\varepsilon \times F \approx p^{-1}(U_\varepsilon)$ is a coordinate function of a coordinate neighborhood $U_\varepsilon \ni b_\varepsilon$, and $\Omega_\varepsilon (= \Omega|_{\varepsilon \times F}): F \approx p^{-1}(b_\varepsilon)$. The homeomorphism $w_\#$ induces an isomorphism $w_\#: \pi_n(F) \cong \pi_n(F)$, and $\pi_n(F)$ forms a \dagger local coefficient on B . Then the cochain $c(f)$ is a \dagger cocycle with the local coefficient $\pi_n(F)$, called the **obstruction cocycle** of f . Furthermore, the set $\{c(f)\}$ for every cross section $f: B^n \rightarrow E$ is a cohomology class $c^{n+1}(\xi) \in H^{n+1}(B; \pi_n(F))$ (local coefficient), and $c^{n+1}(\xi)$ is called the **primary obstruction** to the construction of a cross section. There is a cross section over the $(n+1)$ -skeleton B^{n+1} if and only if $c^{n+1}(\xi) = 0$. The local coefficient $\pi_n(F)$ is trivial if B is \dagger simply connected or, for example, if the structure group G of ξ is connected (in which case ξ is called an **orientable fiber bundle**); when this is true, $c^{n+1}(\xi)$ is an element of $H^{n+1}(B; \pi_n(F))$, where $\pi_n(F)$ is not a local coefficient. Furthermore, if $c^{n+1}(\xi) = 0$ and $\pi_i(F) = 0$ ($n < i < m$), then the secondary obstruction $\Omega^{m+1}(\xi) \in H^{m+1}(B; \pi_m(F))$ is defined similarly (\rightarrow 305 Obstructions).

M. Stiefel-Whitney Classes

Let $\xi = (E, p, B, F, O(n))$ be an $O(n)$ -bundle over an arcwise connected polyhedron B and let ξ^0 be its associated principal bundle. Consider the Stiefel manifold $V_{n,n-k} = V_{n,n-k}(\mathbf{R}) = O(n)/I_{n-k} \times O(k)$, which is $(k-1)$ -connected, and the associated bundle $\xi^k = \xi^0 \times_{O(n)} V_{n,n-k}$ with fiber $V_{n,n-k}$. The primary obstruction $W_{k+1}(\xi) = c^{k+1}(\xi^k) \in H^{k+1}(B; \pi_k(V_{n,n-k}))$ ($k = 0, 1, \dots, n-1$) is called the **Stiefel-Whitney class** of ξ . We have $2W_{k+1}(\xi) = 0$ unless $k = n-1$ and k is odd. Hence we usually consider $W_{k+1}(\xi) \in H^{k+1}(B; \mathbf{Z}_2)$. ξ is orientable, i.e., the group of ξ is reducible to $SO(n)$, if and only if $W_1(\xi) = 0$. The **Stiefel-Whitney classes of an n -dimensional \dagger differentiable manifold M** are defined to be those of the tangent bundle $\mathfrak{T}(M)$. Since the orientability of M coincides with that of $\mathfrak{T}(M)$, M is orientable if and only if $W_1(M) = 0$. The condition $W_{k+1}(M) = 0$ is necessary for the existence of a continuous field of orthonormal tangent $(n-k)$ -frames over M (if $k = n-1$ this condition is also suffi-

cient). Also, when M is closed, $W_n(M)$ is equal to $\chi(M)\mu$, where μ is the fundamental cohomology class of M and $\chi(M)$ is the \dagger Euler characteristic of M (\rightarrow 56 Characteristic Classes B).

N. Chern Classes

For a $U(n)$ -bundle $\xi = (E, p, B, F, U(n))$, the primary obstruction $C_{k+1}(\xi) = c^{2k+2}(\xi^k) \in H^{2k+2}(B; \mathbf{Z})$ ($k=0, 1, \dots, n-1$) of the associated bundle $\xi^k = \xi^0 \times_{U(n)} V_{n,n-k}(\mathbf{C})$ is called the **Chern class** of ξ . If we consider ξ as an $O(2n)$ -bundle by $U(n) \subset O(2n)$, then $W_{2k+1}(\xi) = 0$ and $W_{2k}(\xi) = C_k(\xi) \pmod{2}$. The **Chern classes** of a real $2n$ -dimensional almost complex manifold are defined to be those of the tangent bundle $\mathfrak{T}(M)$ (\rightarrow 56 Characteristic Classes C).

O. Bundles of Class C^r , Analytic Bundles

A fiber bundle $\xi = (E, p, B, F, G)$ is called a **fiber bundle of class C^r** ($r=0, 1, \dots, \infty, \omega$) if E, B, F are differentiable manifolds of class C^r , G is a Lie group and a transformation group of F of class C^r , and p and the coordinate functions are differentiable mappings of class C^r . Bundles of class C^0 are usual G -bundles, and those of class C^ω are **real analytic fiber bundles**. Similarly, **complex analytic fiber bundles** are defined by the notions of complex manifolds, \dagger complex Lie groups, and \dagger holomorphic mappings. For example, the universal bundles $\xi(n-1, O(m))$ and $\eta(2n, U(m))$ are real and complex analytic principal bundles, respectively, and the tangent bundle $\mathfrak{T}(M)$ of a C^{r+1} (or complex) manifold is a C^r (or complex analytic) vector bundle. The operations of the Whitney sum, etc., are defined analogously for these vector bundles.

The equivalence of C^r (complex analytic) bundles is defined by means of bundle mappings that are C^r -differentiable (holomorphic). Bundles of class C^r ($r \leq \infty$) are classified by C^r mappings into a classifying space in the same manner as for bundles of class C^0 . Also, the connection of class C^r (\rightarrow 80 Connections) in C^r bundles is an important notion.

For complex analytic bundles, a similar classification has been obtained for restricted spaces by K. Kodaira, Serre, S. Nakano [8], and others. The classification of complex analytic bundles over a \dagger Stein manifold is reduced to that of bundles of class C^0 (**Oka's principle** [9]), and similar results are valid for C^ω -manifolds. The complex analytic (or holomorphic) connection does not necessarily exist, and M. F. Atiyah [10] found the condition for its existence and its relation with Chern classes.

P. Microbundles

A system $x: B \xrightarrow{i} E \xrightarrow{j} B$ of topological spaces E, B and continuous mappings i, j is called an n -dimensional **microbundle** over B if for each $b \in B$, there exist a neighborhood U of b , a neighborhood V of $i(U)$, and a homeomorphism $h: V \approx U \times \mathbf{R}^n$ with $h \circ i|_U = i_1, j|_V = p_1 \circ h$ ($i_1: U \approx U \times 0 \subset U \times \mathbf{R}^n$, and $p_1: U \times \mathbf{R}^n \rightarrow U$ is the projection). Let $H_0(n)$ be the topological group of all homeomorphisms of \mathbf{R}^n onto itself fixing the origin with compact-open topology. Then the equivalence classes of n -dimensional microbundles over B are naturally in one-to-one correspondence with the equivalence classes of $H_0(n)$ -bundles with base space B and fiber \mathbf{R}^n [13].

In the category of polyhedra and PL (\dagger piecewise linear) mappings, the notion of a **PL microbundle** can be defined in the same manner. The structural group of n -dimensional PL microbundles is defined, in a generalized sense, as a complete \dagger semi-simplicial complex [11]. The **tangent PL microbundle** is defined for any \dagger topological (PL) manifold. J. Milnor classified \dagger smoothings of PL manifolds by means of PL microbundles [11] and then showed that the \dagger tangent vector bundles and its \dagger Pontryagin classes of smooth manifolds are not topologically invariant [12].

For a \dagger PL embedding $f: M \rightarrow N$ between PL manifolds, if there is a neighborhood E of $f(M)$ in N and a PL mapping $p: E \rightarrow M$ so that a diagram $v: M \xrightarrow{f} E \xrightarrow{p} M$ is a PL microbundle, then v is called a **normal PL microbundle** of f . In this case, f is \dagger locally flat.

There is a locally flat PL embedding between PL manifolds which admits no normal PL microbundle [14].

Q. Block Bundles

As the normal bundle theory for locally flat PL embeddings of PL manifolds, the concept of block bundle was introduced independently by C. P. Rourke and B. J. Sanderson [15], M. Kato [16], and C. Morlet [17]. Let E be a \dagger polyhedron, and let K be a cell complex. A set $\{E_\sigma | \sigma \in K\}$ of PL balls in E indexed on K is called a **q -block structure** of E if the following three conditions are satisfied: (1) $\bigcup_{\sigma \in K} E_\sigma = E$; (2) for each $\sigma \in K$, there is a \dagger PL homeomorphism $h_\sigma: \sigma \times I^q \rightarrow E_\sigma$ such that $h_\sigma(\tau \times I^q) = E_\tau$ for each face τ of σ , where $I = [-1, 1]$; and (3) if $E_\sigma \cap E_\rho \neq \emptyset$, then $E_\sigma \cap E_\rho = E_\tau$, where $\tau = \sigma \cap \rho$.

For $\sigma \in K$, E_σ is called the **block** over σ , and h_σ in (2) is called a **trivialization** of E_σ . Then a triple $(E, K, \{E_\sigma | \sigma \in K\})$ is referred to as a **q -block bundle** over K and is denoted by ξ/K . Another block bundle $\xi'/K = (E', K, \{E'_\sigma | \sigma \in$

$K\}$ over the same complex K is said to be **isomorphic** with ξ/K if there is a PL homeomorphism $g: E \rightarrow E'$, called an **isomorphism**, such that $g(E_\sigma) = E'_\sigma$ ($\sigma \in K$). A PL embedding $i: |K| \rightarrow E$ is a **zero-section** of ξ/K if for each $\sigma \in K$, there is a trivialization $h_\sigma: \sigma \times I^q \rightarrow E_\sigma$ such that $h_\sigma(x, 0) = i(x)$ ($x \in \sigma$). In this case we say that ξ/K is a block bundle with a zero-section $i: |K| \rightarrow E$. There is a unique zero-section of ξ/K up to isomorphism of ξ/K onto itself. For every \dagger locally flat PL embedding between PL manifolds M and W of codimension q and for any cell division K of M , a \dagger derived neighborhood N of $f(M)$ in W admits a unique q -block bundle $\nu/K = (N, K, \{N_\sigma | \sigma \in K\})$ with $f: M \rightarrow N$ as a zero-section up to isomorphism respecting the zero-section [15]. The block bundle ν/K is called a **normal block bundle** of $f: M \rightarrow W$.

References

- [1] N. E. Steenrod, The topology of fiber bundles, Princeton Univ. Press, 1951.
- [2] E. Stiefel, Richtungsfelder und Fernparallelismus in n -dimensionalen Mannigfaltigkeiten, Comment. Math. Helv., 8 (1936), 3–51.
- [3] H. Whitney, Topological properties of differentiable manifolds, Bull. Amer. Math. Soc., 43 (1937), 785–805.
- [4] S. S. Chern, Some new view-points in differential geometry in the large, Bull. Amer. Math. Soc., 52 (1946), 1–30.
- [5] F. Hirzebruch, Topological methods in algebraic geometry, Springer, third edition, 1966.
- [6] J. W. Milnor, Construction of universal bundles II, Ann. Math., (2) 63 (1956), 430–436.
- [7] J. C. Moore, Espaces classifiants, Sémin. H. Cartan, 12, no. 5 (1959–1960), Ecole Norm. Sup., Secrétariat Mathématique.
- [8] S. Nakano, On complex analytic vector bundles, J. Math. Soc. Japan, 7 (1955), 1–12.
- [9] H. Grauert, Analytische Faserungen über holomorph-vollständigen Räumen, Math. Ann., 135 (1958), 263–273.
- [10] M. F. Atiyah, Complex analytic connections in fiber bundles, Trans. Amer. Math. Soc., 85 (1957), 181–207.
- [11] J. W. Milnor, Microbundles and differentiable structures, Lecture notes, Princeton Univ., 1961.
- [12] J. W. Milnor, Microbundles I, Topology, 3 (1964), Suppl. 1, 53–80.
- [13] J. Kister, Microbundles are fiber bundles, Bull. Amer. Math. Soc., 69 (1963), 854–857.
- [14] C. P. Rourke and B. J. Sanderson, An imbedding without a normal bundle, Inventiones Math., 3 (1967), 293–299.
- [15] C. P. Rourke and B. J. Sanderson, Block-bundles I–III, Ann. Math., (2) 87 (1968).
- [16] M. Kato, Combinatorial prebundles I, II, Osaka J. Math., 4 (1967).
- [17] C. Morlet, Les méthodes de la topologie différentielle dans l'étude des variétés semi-lineaires, Ann. Sci. École Norm. Sup., 1 (1968), 313–394.
- [18] D. Husemoller, Fiber bundles, McGraw-Hill, 1966.

148 (IX.10) Fiber Spaces

A. General Remarks

J.-P. Serre [1] generalized the concept of fiber bundles to that of fiber spaces by utilizing the covering homotopy property (\rightarrow Section B). He applied the theory of \dagger spectral sequences, due to J. Leray, to the (cubic) \dagger singular (co)-homology groups of fiber spaces. These are quite useful for determining (co)homology structures and homotopy groups of topological spaces, and are now of fundamental importance in algebraic topology.

B. Definitions

Let $p: E \rightarrow B$ be a continuous mapping between topological spaces, and let X be a topological space. Then we say that p has the **covering homotopy property** with respect to X if for any mapping $f: X \rightarrow E$ and \dagger homotopy $g: X \rightarrow B$ ($0 \leq t \leq 1$) with $p \circ f = g_0$, there is a homotopy $f_t: X \rightarrow E$ ($0 \leq t \leq 1$) with $f_0 = f$ and $p \circ f_t = g_t$. We call (E, p, B) a **fiber space** (or **fibration**) if p has the covering homotopy property with respect to every cube $I^n = \{(x_1, \dots, x_n) | 0 \leq x_i \leq 1\}$, $n = 0, 1, \dots$ (then p has the covering homotopy property with respect to every \dagger CW complex). Then E is called the **total space**, p the **projection**, B the **base space**, and $F_b = p^{-1}(b)$ the **fiber** over $b \in B$.

Let E, B, F be topological spaces and $p: E \rightarrow B$ a continuous mapping. We call (E, p, B, F) a **locally trivial fiber space** if for each $b \in B$, there exist an open neighborhood U of b and a homeomorphism $\varphi: U \times F \rightarrow p^{-1}(U)$ with $p \circ \varphi(b', y) = b'$ ($b' \in U, y \in F$). In this case, p has the covering homotopy property with respect to each \dagger paracompact space; hence (E, p, B) is a fiber space. A \dagger fiber bundle is clearly a locally trivial fiber space.

C. Path Spaces

Another important example of a fiber space is a path space. A **path** in a topological space X is a continuous mapping $w: I \rightarrow X$ ($I = [0, 1]$). Given subsets A_0 and A_1 of X , the **path space** $\Omega(X; A_0, A_1)$ is the space of all paths $w: I \rightarrow X$ satisfying $w(\varepsilon) \in A_\varepsilon$ ($\varepsilon = 0, 1$) topologized by \dagger compact-open topology. Define $p_\varepsilon: \Omega(X; A_0, A_1) \rightarrow A_\varepsilon$ by $p_\varepsilon(w) = w(\varepsilon)$ ($\varepsilon = 0, 1$). Then $(\Omega(X; A_0, A_1), p_\varepsilon, A_\varepsilon)$ is a fiber space; in fact, p_ε has the covering homotopy property with respect to every topological space. In particular, the total space of the fiber space $(\Omega(X; X, *), p_0, X)$ ($* \in X$) is \dagger contractible, and the fiber $p_0^{-1}(*) = \Omega(X; *, *) = \Omega X$ over $*$ is the \dagger loop space of X with base point $*$. For a continuous mapping $f: Y \rightarrow X$, consider the space $E_f = \{(y, w) \in Y \times \Omega(X; X, X) \mid f(y) = w(0)\}$ and the continuous mapping $p: E_f \rightarrow X$ defined by $p(y, w) = w(1)$. Then $Y \subset E_f$, and Y is a \dagger deformation retract of E_f ; furthermore, (E_f, p, X) is a fiber space with $f = p \mid Y$.

D. Homotopy Groups of Fiber Spaces

For \dagger homotopy groups of fiber spaces, the **Hurewicz-Steenrod isomorphism theorem** holds: Let (E, p, B) be a fiber space and $F = p^{-1}(*)$ the fiber over the base point $* \in B$. Then $p_*: \pi_n(E, F) \rightarrow \pi_n(B)$ is an isomorphism for $n \geq 2$ and a bijection for $n = 1$. By this theorem, we have the **homotopy exact sequence** of a fiber space:

$$\dots \rightarrow \pi_{n+1}(B) \xrightarrow{\Delta} \pi_n(F) \xrightarrow{i_*} \pi_n(E) \xrightarrow{p_*} \pi_n(B) \rightarrow \dots$$

Furthermore, the more general exact sequence

$$\dots \rightarrow \pi(Z; \Omega B)_0 \rightarrow \pi(Z; F)_0 \xrightarrow{i_*} \pi(Z; E)_0 \xrightarrow{p_*} \pi(Z; B)_0$$

is valid for each CW complex Z , where $\pi(Z; X)_0$ denotes the \dagger homotopy set of mappings from Z to X relative to the base point.

Example (1). A **cross section** of a fiber space (E, p, B) is a continuous mapping $f: B \rightarrow E$ with $p \circ f = 1$. If (E, p, B) has a \dagger cross section or the fiber F is a \dagger retract of E , then $\pi_n(E) \cong \pi_n(B) + \pi_n(F)$. If F is contractible in E , then $\pi_n(B) \cong \pi_n(E) + \pi_{n-1}(F)$ ($n \geq 2$).

Example (2). (E, p, B) is called an **n -connective fiber space** if B is \dagger arcwise connected, E is $\dagger n$ -connected, and $p_*: \pi_i(E) \rightarrow \pi_i(B)$ is an isomorphism, for $i > n$. For each arcwise connected space B and integer n , there is such a fiber space.

Example (3). For a CW complex X , there are topological spaces X_n and continuous mappings $f_n: X \rightarrow X_n$, $q_{n+1}: X_{n+1} \rightarrow X_n$ ($n = 0, 1, \dots$) with the following four properties: (i) X_n ($0 \leq n \leq m$) is a point if X is m -connected; (ii) $f_n^*: \pi_i(X) \cong \pi_i(X_n)$ ($i \leq n$); (iii) (X_n, q_n, X_{n-1}) is a fiber

space, and its fiber is an \dagger Eilenberg-MacLane space $K(\pi_n(X), n)$; (iv) $q_n \circ f_n$ is \dagger homotopic to f_{n-1} . Such a system $\{X_n, f_n, q_n\}$ is called the **Postnikov system** of X and is in a sense considered a decomposition of X into Eilenberg-MacLane spaces.

E. Spectral Sequences of Fiber Spaces

The cohomological properties of fiber spaces are obtained mainly from the following results (which are valid similarly for homology except for properties of products). Assume that the base space B of a given fiber space (E, p, B) is \dagger simply connected and the fiber $F = p^{-1}(*)$ is arcwise connected, and let R be a commutative ring with unit. Then the **spectral sequence** (of singular cohomology) **of the fiber space** (E, p, B) (with coefficients in R) is defined to be a sequence $\{E_r, d_r\}$ satisfying the following properties:

- (i) $E_r = \sum_{p,q} E_r^{p,q}$ are bigraded R -modules, and $d_r = \sum_{p,q} d_r^{p,q}$ are R -linear differentials such that $d_r(E_r^{p,q}) \subset E_r^{p+r, q-r+1}$.
- (ii) $E_{r+1}^{p,q} = \text{Ker } d_r^{p,q} / \text{Im } d_r^{p-r, q+r-1}$, which means that $E_{r+1} = H(E_r)$.
- (iii) $E_r^{p,q} = 0$ for $p < 0$ or $q < 0$, $E_r^{p,q} = E_{r+1}^{p,q} = \dots = E_\infty^{p,q}$ for $r > \max(p, q+1)$.
- (iv) E_r has a product for which $E_r^{p,q} \cdot E_r^{p',q'} \subset E_r^{p+p', q+q'}$ and $d_r(u \cdot v) = (d_r u) \cdot v + (-1)^{p+q} u \cdot d_r v$ ($u \in E_r^{p,q}$). Furthermore, the induced product in $H(E_r)$ coincides with the product in E_{r+1} .
- (v) E_∞ is the \dagger bigraded module associated with some filtration of the cohomology module $H^*(E; R)$; that is, $H^n(E; R) = D^{0,n} \supset D^{1,n-1} \supset \dots \supset D^{n,0} \supset D^{n+1,-1} = 0$ and $E_\infty^{p,q} = D^{p,q} / D^{p+1, q-1}$. Furthermore, the \dagger cup product \smile in $H^*(E; R)$ satisfies $D^{p,q} \smile D^{p',q'} \subset D^{p+p', q+q'}$ and coincides with the given product in E_∞ .
- (vi) $E_2^{p,q} = H^p(B; H^q(F; R))$, and the product in E_2 coincides with the cup product in $H^*(B; H^*(F; R))$.
- (vii) The composition of $H^n(B; R) = E_2^{n,0} \rightarrow E_3^{n,0} \rightarrow \dots \rightarrow E_{n+1}^{n,0} = E_\infty^{n,0} = D^{n,0} \subset H^n(E; R)$ is equal to p^* , and the composition of $H^n(E; R) = D^{0,n} \rightarrow E_\infty^{0,n} = E_{n+2}^{0,n} \subset \dots \subset E_3^{0,n} \subset E_2^{0,n} = H^n(F; R)$ is equal to i^* ($i: F \subset E$), where each \rightarrow is the projection onto the quotient module. In the sequence

$$H^{n-1}(F; R) \xrightarrow{\partial^*} H^n(E; R) \xleftarrow{p^*} H^n(B; R),$$

we have $\partial^{*-1}(\text{Im } p^*) = E_\infty^{0, n-1}$, $\text{Coim } p^* = E_\infty^{n,0}$, and $d_n: E_\infty^{0, n-1} \rightarrow E_\infty^{n,0}$ is equal to the **transgression** $\tau^* = p^{*-1} \circ \partial^*: \partial^{*-1}(\text{Im } p^*) \rightarrow \text{Coim } p^*$. Each element of $\partial^{*-1}(\text{Im } p^*)$ is called **transgressive**.

In the following examples, we assume that R is a principal ideal ring.

Example (4). Let k be a commutative

field, and assume that $\dim_k H_*(B; k)$ and $\dim_k H_*(F; k)$ are finite. Then for the \dagger Poincaré polynomial $P_X(t) = \sum_n b_n t^n$, $b_n = \dim_k H_n(X; k)$, we have $P_E(t) = P_B(t)P_F(t) - (1+t)\varphi(t)$, where $\varphi(t)$ is a polynomial with nonnegative coefficients (Leray). In particular, for the \dagger Euler characteristic $\chi(X) = P_X(-1)$, we have $\chi(E) = \chi(B)\chi(F)$. Also, if $i_*: H_n(F; k) \rightarrow H_n(E; k)$ is monomorphic for each $n \geq 0$, then $P_E(t) = P_B(t)P_F(t)$.

Example (5). Isomorphism theorem: If $H_n(B; R) = 0$ ($0 < n < r$) and $H_n(F; R) = 0$ ($0 < n < s$), then $p_*: H_n(E, F; R) \rightarrow H_n(B; R)$ is isomorphic for $0 < n < r + s$ and epimorphic for $n = r + s$, and we have the following **homology exact sequence**:

$$\begin{aligned} \dots \rightarrow H_n(F; R) \xrightarrow{i_*} H_n(E; R) \xrightarrow{p_*} H_n(B; R) \\ \xrightarrow{\tau_*} H_{n-1}(F; R) \rightarrow \dots, \quad n < r + s \end{aligned}$$

(similarly for the cohomology).

Example (6). Assume that $H^n(F; R) \cong H^n(S^r; R)$ (S^r is the r -sphere, $r \geq 1$). Let $\tilde{E} = M_p$ be the \dagger mapping cylinder of $p: E \rightarrow B$ and $\tilde{p}: \tilde{E} \rightarrow B$ be the continuous mapping defined by p . Then the **Thom-Gysin isomorphism** $\tilde{g}: H^{n-r-1}(B; R) \cong H^n(\tilde{E}, E; R)$ ($n \geq 0$) with $\tilde{g}(\alpha) = \tilde{p}^*(\alpha) \smile \tilde{g}(1)$ holds (\rightarrow 114 Differential Topology G). Also, we have the **Gysin exact sequence**:

$$\begin{aligned} \dots \rightarrow H^n(B; R) \xrightarrow{p^*} H^n(E; R) \rightarrow H^{n-r}(B; R) \\ \xrightarrow{g} H^{n+1}(B; R) \rightarrow \dots, \end{aligned}$$

where g satisfies $g(\alpha) = \alpha \smile \Omega = \Omega \smile \alpha$ ($\Omega = g(1) \in H^{r+1}(B; R)$). Here Ω is equal to the image of a generator of $H^r(F; R) = R$ by the transgression τ^* , and $2\Omega = 0$ if r is even. (These results hold also for $r = 0$ and $R = \mathbb{Z}_2$.)

Example (7). Assume that $H^n(B; R) \cong H^n(S^r; R)$ ($r \geq 2$). Then we have $H^{n-r}(F; R) \cong H^n(E, F; R)$ ($n \geq 0$) and the **Wang exact sequence**:

$$\begin{aligned} \dots \rightarrow H^n(E; R) \xrightarrow{i^*} H^n(F; R) \\ \xrightarrow{\theta} H^{n-r+1}(F; R) \rightarrow H^{n+1}(E; R) \rightarrow \dots, \end{aligned}$$

where θ satisfies $\theta(\alpha \smile \beta) = \theta(\alpha) \smile \beta + (-1)^{n(r-1)} \alpha \smile \theta(\beta)$ ($\alpha, \beta \in H^n(F; R)$).

Example (8). For a field k of odd characteristic, if $H^n(E; k) = 0$ ($n > 0$) and the algebra $H^*(F; k)$ is generated by a finite number of elements of odd degree, then $H^*(F; k) \cong \wedge_k(x_1, \dots, x_l)$ (the \dagger exterior algebra) and $H^*(B; k) \cong k[y_1, \dots, y_l]$, where $y_i = \tau^*(x_i)$ (A. Borel).

References

[1] J.-P. Serre, Homologie singulière des espaces fibrés, Ann. Math., (2) 54 (1951), 425–505.

[2] S. T. Hu, Homotopy theory, Academic Press, 1959.

[3] A. Borel, Sur la cohomologie des espaces fibrés principaux et des espaces homogènes de groupes de Lie compacts, Ann. Math., (2) 57 (1953), 115–207.

[4] E. H. Spanier, Algebraic topology, McGraw-Hill, 1966.

[5] G. W. Whitehead, Elements of homotopy theory, Springer, 1978.

149 (III.6) Fields

A. Definition

A set K having at least two elements is called a **field** if two operations, called \dagger addition ($+$) and \dagger multiplication (\cdot), are defined in K and satisfy the following three axioms.

(1) For any two elements a, b of K , the sum $a + b$ is defined; the associative law $(a + b) + c = a + (b + c)$ and the commutative law $a + b = b + a$ hold; and there exists for arbitrary a, b a unique element x such that $a + x = b$; that is, K is an \dagger Abelian group with respect to the addition (the \dagger identity element of this group is denoted by 0 and is called the **zero element** of K).

(2) For any two elements a, b of K , the product ab ($= a \cdot b$) is defined; the associative law $(ab)c = a(bc)$ and the commutative law $ab = ba$ hold; and there exists for arbitrary a, b with $a \neq 0$ a unique element x such that $ax = b$, that is, the set K^* of all nonzero elements of K is an Abelian group with respect to multiplication. K^* is called the **multiplicative group** of K , while the identity element of K^* is denoted by 1 and is called the **unity element** (**unit element** or **identity element**) of K .

(3) The distributive law $a(b + c) = ab + ac$ holds. In other words, a field is a \dagger commutative ring whose nonzero elements form a group with respect to multiplication.

A noncommutative ring whose nonzero elements form a group is called a **noncommutative field** (**skew field** or **s-field**). It should be noted that sometimes a field is defined as a ring whose nonzero elements form a group without assuming the commutativity of that group, and in this case our “field” defined before is called a commutative field. (The term “skew field” is sometimes used to mean either a commutative or a noncommutative field.) In this article we limit ourselves to commutative fields (for noncommutative fields \rightarrow 29 Associative Algebras).

B. General Properties

Since a field K is a commutative ring, we have properties such as $a0 = 0a = 0$, $(-a)b = a(-b) = -ab$ for elements a, b in K . If a †subring k of K is a field, we say that k is a **subfield** of K or K is an **overfield** (**extension field** or simply **extension**) of k . If a field K has no subfield other than K , K is called a **prime field**.

A map f of a field K into another field K' is called a (field) **homomorphism** if it is a ring homomorphism, i.e., if it satisfies $f(a + b) = f(a) + f(b)$, $f(ab) = f(a)f(b)$. Since a field is †simple as a ring, every (field) homomorphism is an injection unless it maps everything to zero. A homomorphism of K into K' is called an **isomorphism** if it is a bijection, and K and K' are called isomorphic if there exists an isomorphism of K onto K' . An isomorphism of K onto itself is called an **automorphism** of K .

If there is a natural number n such that the sum $n1 = \overbrace{1 + \dots + 1}^n$ of the unity element 1 is 0 , then the minimum of such n is a prime number p , called the **characteristic** of K . On the other hand, if there is no natural number n such that $n1 = 0$, we say that the characteristic of K is 0 .

C. Examples of Fields

The rational number field \mathbf{Q} consisting of all rational numbers, the real number field \mathbf{R} consisting of all real numbers, and the complex number field \mathbf{C} consisting of all complex numbers, are all fields of characteristic 0 . A subfield of the complex number field \mathbf{C} is called a **number field**. The rational number field is a prime field, and every prime field of characteristic 0 is isomorphic to the rational number field. For the ring \mathbf{Z} of all rational integers, the residue class ring modulo a prime number p is a field $\mathbf{Z}/p\mathbf{Z} = \{0, 1, 2, \dots, p-1 \pmod{p}\}$ of characteristic p , called the **residue class field** for p . Thus $\mathbf{Z}/p\mathbf{Z}$ is a prime field, and every prime field of characteristic p is isomorphic to $\mathbf{Z}/p\mathbf{Z}$. If the number of the elements of a field K is finite, K is called a **finite field**. $\mathbf{Z}/p\mathbf{Z}$ is an example of a finite field.

D. Extensions of a Field

In order to express that K is an extension field of k , we often use the notation K/k . Subfields of K containing k are called **intermediate fields** of K/k . Consider two extensions K_1/k_1 and K_2/k_2 , and let $\varphi: K_1 \rightarrow K_2$ be an isomorphism which induces an isomorphism $\psi: k_1 \rightarrow k_2$. Then we call φ an **extension** of ψ . Suppose that k_1, K_2 are given fields and K_2 contains a subfield k_2 isomorphic to k_1 . Then there exist

an extension field K_1 of k_1 and an isomorphism $\varphi: K_1 \rightarrow K_2$ which is an extension of the given isomorphism $\psi: k_1 \rightarrow k_2$; construction of the field K_1 is often called the embedding of k_1 into K_2 . When K_1 and K_2 are extensions of k , an isomorphism $K_1 \rightarrow K_2$ is called a ***k*-isomorphism** if it leaves every element of k invariant.

In an extension K/k , let S be a subset of K . The smallest intermediate field of K/k containing S is called the field obtained by **adjoining** S to k or the field **generated** by S over k , denoted by $k(S)$. The field $k(S)$ consists of those elements in K each of which is a rational expression in a finite number of elements of S with coefficients in k . An extension field $k(t)$ obtained by adjoining a single element t to k is called a **simple extension** of k , and in this case t is called a **primitive element** of the extension. The †rational function field $k(X)$ with coefficient field k is a simple extension of k with a primitive element X .

When subfields k_λ ($\lambda \in \Lambda$) of a field K are given, the smallest subfield of K containing all these subfields exists and is called the **composite field** of the k_λ .

E. Algebraic and Transcendental Extensions

An element α of an extension field K of a field k is called an **algebraic element** over k if α is a †zero point of a nonzero polynomial, say, $f(X) = a_0 + a_1X + \dots + a_nX^n$ with coefficients in k . If α is not algebraic over k , then α is called a **transcendental element** over k . An algebraic element α is always a root of an irreducible polynomial over k which is uniquely determined up to a constant factor ($\in k^*$) and is called the **minimal polynomial** of α over k . K is called an **algebraic extension** of k if all elements of K are algebraic over k ; otherwise we call K a **transcendental extension** of k . If K_1 is an algebraic extension of K and K is an algebraic extension of k , then K_1 is also an algebraic extension of k . In an arbitrary extension field K of k , the set of all algebraic elements over k forms an algebraic extension field of k . A simple extension $k(t)$ with a transcendental element t is isomorphic to the rational function field of one variable with coefficient field k . If t is an algebraic element over k , $k(t)$ is isomorphic to the †residue class field of the polynomial ring $k[X]$ modulo the minimal polynomial $f(X)$ of t over k .

F. Finite Extensions

An extension field K of a field k is called a **finite extension** if K has no infinite set of elements that are †linearly independent over k ,

i.e., if K is a finite-dimensional linear space over k . The dimension of the linear space over k is called the **degree** of K over k and is denoted by $(K:k)$ (or $[K:k]$). If K is a finite extension of k and L is a finite extension of K , then L is also a finite extension of k and $(L:K)(K:k) = (L:k)$. Every finite extension field of k is an algebraic extension of k and is obtained by adjoining a finite number of algebraic elements to k . Conversely, every field obtained by adjoining a finite number of algebraic elements to k is a finite extension of k . If $K = k(\alpha)$ with an algebraic element α , then $(K:k)$ is equal to the degree of the minimal polynomial of α over k , also called the **degree** of α over k . Every element of $k(\alpha)$ is expressed as a polynomial in α with coefficients in k . On the other hand, for any nonconstant polynomial $f(X)$ of $k[X]$ there exists a simple extension $k(\alpha)$ such that α is a root of $f(X)$.

G. Normal Extensions

An algebraic extension field K of a field k is called a **normal extension** of k if every irreducible polynomial of $k[X]$ which has a root in K can always be decomposed into a product of linear factors in $K[X]$. An extension field K of k is called a **splitting field** of a (nonconstant) polynomial $f(X) \in k[X]$ if $f(X)$ can be decomposed as a product of linear polynomials, i.e., $f(X) = c(X - \alpha_1)(X - \alpha_2) \dots (X - \alpha_n)$, $c \in k$, $\alpha_i \in K$. A splitting field K of $f(X)$ ($f(X) \in k[X]$) is called a **minimal splitting field** of $f(X)$ if any proper subfield L of K ($K \supset L \supset k$) is not a splitting field of $f(X)$. A minimal splitting field of $f(X)$ is obtained by adjoining all the zero points of $f(X)$. A finite extension field of k is a normal extension if and only if it is a minimal splitting field of a polynomial of $k[X]$. For any given (nonconstant) polynomial $f(X) \in k[X]$, there exists a minimal splitting field of $f(X)$, and all minimal splitting fields of $f(X)$ are k -isomorphic.

H. Separable and Inseparable Extensions

An algebraic element α over k is called a **separable element** or an **inseparable element** over k according as the minimal polynomial of α over k is 'separable' or 'inseparable'. An algebraic extension K of k is called a **separable extension** of k if all the elements of K are separable over k ; otherwise, K is called an **inseparable extension**. An element α is separable with respect to k if and only if the minimal polynomial of α over k has no double root in its splitting field. If α is inseparable, then k has nonzero characteristic p , and the minimal polynomial $f(X)$ of α can be decomposed as $f(X) =$

$(X - \alpha_1)^{p^r}(X - \alpha_2)^{p^r} \dots (X - \alpha_m)^{p^r}$, $r \geq 1$, where $\alpha_1, \alpha_2, \dots, \alpha_m$ are distinct roots of $f(X)$ in its splitting field; between the degree n of $f(X)$ and the number m of distinct roots of $f(X)$, the relation $n = mp^r$ holds. In particular, if $\alpha^{p^r} \in k$ for some r , we call α a **purely inseparable element** over k . An algebraic extension of k is called **purely inseparable** if all elements of the field are purely inseparable over k . In an algebraic extension K of k the set of all separable elements forms an intermediate field K_0 of K/k . The field K_0 is called the **maximal separable extension** of k in K . If K is inseparable over k , i.e., if $K \neq K_0$, then the characteristic of k is $p \neq 0$, and K is purely inseparable over K_0 . The degrees $d = [K_0:k]$ and $p^r = [K:K_0]$ are denoted by $[K:k]_s$ and $[K:k]_i$, respectively. A separable extension of a separable extension of k is also separable over k , and every finite separable extension of k is a simple extension.

If no inseparable irreducible polynomial in $k[X]$ exists, we call k a **perfect field**; otherwise, an **imperfect field**. Every field of characteristic 0 is a perfect field. A field of characteristic p ($\neq 0$) is perfect if and only if for each $a \in k$ the polynomial $X^p - a$ has a root in k . Every algebraic extension of a perfect field is a separable extension and a perfect field. Any imperfect field has an inseparable, in fact purely inseparable, proper extension.

I. Algebraically Closed Fields

If every nonconstant polynomial of $k[X]$ can be decomposed into a product of linear polynomials of $k[X]$, or equivalently, if every irreducible polynomial of $k[X]$ is linear, k is called an **algebraically closed field**; k is algebraically closed if and only if k has no algebraic extension field other than k , and hence every algebraically closed field is perfect. For any given field k there exists an algebraically closed algebraic extension field of k unique up to k -isomorphisms (E. Steinitz); hence we call such a field the **algebraic closure** of k . To proceed further, suppose that we are given a field k and its extension K . If there is no algebraic element of K over k outside of k , i.e., if k is the intersection of K and the algebraic closure of k , then we say that k is **algebraically closed** in K . The complex number field is an algebraically closed field (C. F. Gauss's fundamental theorem of algebra; \rightarrow 10 Algebraic Equations).

J. Conjugates

Let k be a field and K an algebraic extension of k . Two elements α, β of K are called **conju-**

gate over k if they are roots of the same irreducible polynomial of $k[X]$ (or equivalently, if the minimal polynomials of α and β with respect to k coincide); in this case we call the subfields $k(\alpha)$, $k(\beta)$ **conjugate fields** over k . The conjugate fields $k(\alpha)$ and $k(\beta)$ are k -isomorphic under an isomorphism σ such that $\sigma(\alpha) = \beta$. In particular, if K is a normal extension of k , the number of conjugate elements of an element α of K is the number of distinct roots of the minimal polynomial $f(X)$ of α , which is independent of the choice of a normal extension K containing k . The element α is separable if and only if the number of conjugate elements in K is the same as the degree of $f(X)$. On the other hand, $k(\alpha)$ is normal over k if and only if $k(\alpha)$ coincides with all its conjugate fields.

Let α be a separable algebraic element over k , and let $\alpha_1 = \alpha, \alpha_2, \dots, \alpha_n$ be conjugate elements of α over k . The product $A = \alpha_1 \alpha_2 \dots \alpha_n$ and sum $B = \alpha_1 + \alpha_2 + \dots + \alpha_n$ are elements of k . Indeed, if $f(X) = X^n + c_1 X^{n-1} + \dots + c_n$ is the minimal polynomial of α with respect to k , we have $A = (-1)^n c_n$, $B = -c_1$, and A and B are called the **norm** and the **trace** of α , respectively, denoted by $A = N(\alpha)$, $B = Tr(\alpha)$. Let K be a finite separable extension of degree n over k , and let α be an element of K . Then the degree m of the minimal polynomial of α is a divisor of n ; that is, $n = mr$ with a positive integer r . We define the norm and the trace of α with respect to K/k by $N_{K/k}(\alpha) = N(\alpha)^r$, $Tr_{K/k}(\alpha) = rTr(\alpha)$, respectively. Then these quantities satisfy $N_{K/k}(\alpha\beta) = N_{K/k}(\alpha)N_{K/k}(\beta)$, $Tr_{K/k}(\alpha + \beta) = Tr_{K/k}(\alpha) + Tr_{K/k}(\beta)$ for $\alpha, \beta \in K$. (For the Galois theory of algebraic extensions \rightarrow 172 Galois Theory.)

K. Transcendental Extensions

Let K be an extension of k and u_1, \dots, u_n be elements of K . An element v of K is said to be **algebraically dependent** on the elements u_1, u_2, \dots, u_n if v is algebraic over the field $k(u_1, u_2, \dots, u_n)$. A subset S of K is called **algebraically independent** over k if no $u \in S$ is algebraically dependent on a finite number of elements of S different from u itself; S is called a **transcendence basis** of K over k if S is algebraically independent and K is algebraic over $k(S)$. Furthermore, if K is separable over $k(S)$, then S is called a **separating transcendence basis** of K over k . There always exists an algebraically independent basis of K over k , and the cardinal number of S depends only on K and k ; this cardinal number is called the **transcendence degree** (or **degree of transcendency**) of K over k . (When S is an infinite set, we sometimes say that the transcendence degree is

infinite.) In particular, if $K = k(S)$ with an algebraically independent S , K is called a **purely transcendental extension** of k .

An extension K of k is called a **separably generated extension**, or simply a **separable extension**, if every finitely generated intermediate field of K/k has a separating transcendence basis over k . If K itself has a separating transcendence basis over k , then K is separably generated, but not conversely.

A purely transcendental extension field of k having a finite transcendence degree n is also called a **rational function field in n variables** over k , and a finite extension of such a rational function field is called an **algebraic function field in n variables** over k .

Let K and L be extension fields of k , both contained in a common extension field. We say that K and L are **linearly disjoint** over k if every subset of K linearly independent over k is also linearly independent over L , or equivalently, if every subset of L linearly independent over k is also linearly independent over K . An algebraic function field $K = k(x_1, x_2, \dots, x_n)$ over k (whose transcendence degree is $\leq n$) is called a **regular extension** of k if K and the algebraic closure \bar{k} of k are linearly disjoint. In order that K be regular over k it is necessary and sufficient that k be algebraically closed in K and that K be separably generated over k .

L. Derivations

A map D of a field K into itself is called a **derivation** of K if it satisfies $D(a+b) = D(a) + D(b)$ and $D(ab) = aD(b) + bD(a)$ for all $a, b \in K$. The set of elements c of K for which $D(c) = 0$ is a subfield. If the characteristic of K is $p (\neq 0)$, then $D(x^p) = 0$ for all $x \in K$. Let k be a subfield of K . A derivation D of K is called a **derivation over k** if $D(c) = 0$ for all $c \in k$; the totality of derivations over k is a ${}^k k$ -module. If $K = k(x_1, x_2, \dots, x_n)$ is an algebraic function field over k , then the k -module of derivations over k has finite dimension $s (\leq n)$ and we can choose s suitable elements u_1, u_2, \dots, u_s of K such that K is separably algebraic over $k(u_1, u_2, \dots, u_s)$. Generally, the transcendence degree r of K over k does not exceed s , and K is separably generated over k if and only if $r = s$.

M. Finite Fields

Finite fields were first considered by E. Galois (1830), so they are also called **Galois fields**. There is no finite noncommutative field (**Wedderburn's theorem**, J. H. M. Wedderburn, *Trans. Amer. Math. Soc.*, 6 (1905)). A simple

proof for this was given by E. Witt (*Abh. Math. Sem. Univ. Hamburg*, 8 (1931)). The characteristic of a finite field is a prime p , and the number of elements of the field is a power of p . Conversely, for any given prime number p and natural number α , there exists a finite field with p^α elements. Such a field is unique up to isomorphism, which we denote by $GF(p^\alpha)$ or $F_q(q=p^\alpha)$. For any positive integer m , $GF(p^{m\alpha})$ is an extension field of $GF(p^\alpha)$ of degree m and a \dagger cyclic extension. Every element a of $GF(p^\alpha)$ satisfies $a^{p^\alpha} = a$; hence a has its p th root in $GF(p^\alpha)$. Therefore every finite field is perfect. The multiplicative group of $GF(p^\alpha)$ is a cyclic group of order $p^\alpha - 1$.

N. Ordered Fields and Real Fields

A field K is called an **ordered field** if there is given a \dagger total order in K such that $a > b$ implies $a + c > b + c$ for all c and $a > b, c > 0$ implies $ac > bc$. The characteristic of an ordered field is always 0. An element a of K is called a **positive element** or a **negative element** according as $a > 0$ or $a < 0$. For an element a of K the **absolute value** of a , denoted by $|a|$, is a or $-a$ according as $a \geq 0$ or $a \leq 0$. If we define neighborhoods of a by the sets $\{x | a - \varepsilon < x < a + \varepsilon\}$ with positive elements ε , K becomes a \dagger Hausdorff space. If, for any two positive elements a, b of K , there exists a natural number n such that $na > b$, then we call K an **Archimedean ordered field**. Two ordered fields are called **similarly isomorphic** if there exists an isomorphism between them under which positive elements are always mapped to positive elements. The rational number field and the real number field are examples of Archimedean ordered fields, while every Archimedean ordered field is similarly isomorphic to a subfield of the real number field. (For the structure of non-Archimedean ordered fields, see [8].) A field k is called a **formally real field** (or simply **real field**) if -1 (1 is the unity element of k) cannot be expressed as a finite sum of squares of elements of k . The real number field is a model of formally real fields. More generally, every ordered field is a formally real field. A formally real field is called a **real closed field** if no proper algebraic extension of it is a formally real field. The real number field is a real closed field. The algebraic closure of a real closed field is obtained by adjoining a root of the polynomial $X^2 + 1$. If a is a nonzero element of a real closed field, then either a or $-a$ can be a square of an element of the field. Every real closed field can be made an ordered field in a unique way, namely, by defining squares of nonzero elements to be positive

150 A Field Theory

elements. Since it is known that every formally real field is a subfield of a real closed field, it follows that every formally real field is an ordered field and therefore is of characteristic 0. The problem of constructing an ordered field out of a formally real field is closely related to the existence of valuations of a certain type [8]. The notion of formally real fields was introduced by E. Artin (*Abh. Math. Sem. Univ. Hamburg*, 5 (1927)). By making use of the theory of formally real fields, Artin succeeded in solving affirmatively \dagger Hilbert's 17th problem, which asked whether every positive definite rational expression (i.e., a rational expression with real coefficients that takes positive values for all real variables) can be expressed as a sum of squares of rational expressions. More precisely, it was shown by A. Pfister that every positive definite function in $\mathbf{R}(X_1, \dots, X_n)$ is a sum of at most 2^n squares.

References

- [1] E. Steinitz, *Algebraische Theorie der Körper*, J. Reine Angew. Math., 137 (1910), 167–309.
- [2] E. Steinitz, *Algebraische Theorie der Körper*, Walter de Gruyter, 1930 (Chelsea, 1950).
- [3] H. Hasse, *Höhere Algebra I, II*, Sammlung Götschen, Walter de Gruyter, 1926–1927.
- [4] B. L. van der Waerden, *Algebra I*, Springer, seventh edition, 1966.
- [5] A. A. Albert, *Fundamental concepts of higher algebra*, Univ. of Chicago Press, 1956.
- [6] N. Bourbaki, *Éléments de mathématique*, *Algèbre*, ch. 5, *Actualités Sci. Ind.*, 1102b, Hermann, second edition, 1959.
- [7] N. Jacobson, *Lectures in abstract algebra III*, Van Nostrand, 1964.
- [8] M. Nagata, *Field theory*, Dekker, 1977.

150 (XX.28) Field Theory

A. History

When a quantity $\psi(x)$, such as velocity, is defined at every point x in a certain region of a space, we say that a field of the quantity ψ is given. This general concept is used in many branches of science. Here we confine ourselves to some branches of physics, in particular to the quantum theory of fields, which describes \dagger elementary particles.

The \dagger theories of elasticity and \dagger hydrodynamics (in particular, concerning \dagger Euler's

equation of motion) deal with displacement and velocity fields, respectively. However, a field in a vacuum (ether), which is quite different from a field in a space filled with matter, first became a subject of physics in [†]electromagnetism. M. Faraday (1837) introduced the electromagnetic field and discovered its fundamental laws, and J. C. Maxwell (1837) completed the mathematical formulation. On the basis of this formalism, A. Einstein (1905) established the theory of [†]relativity and later developed the general theories of relativity and of gravity.

Although quantum theory originated from the problem of blackbody radiation, the quantum theory of the electromagnetic field was first developed by P. A. M. Dirac (1927) after the development of [†]quantum mechanics. Along similar lines P. Jordan and E. P. Wigner (1927) quantized the matter wave (electron field), and W. Heisenberg and W. Pauli (1929) developed the quantum theory of wave fields in general. Subsequently, Jordan and Pauli (1927), Pauli (1939), S. Tomonaga (1943), J. Schwinger (1948), and others reformulated the theory in a relativistically covariant manner. Quantum electrodynamics, dealing with an electromagnetic field interacting with electrons (and positrons), has given excellent agreement with experimental measurements when formulated in this way. Divergence difficulties inherent to quantized field theory were bypassed by the renormalization procedure of Tomonaga, Schwinger, R. P. Feynman, and F. J. Dyson (1947).

On the other hand, H. Yukawa (1934) applied the concept of the quantized field to the interpretation of nuclear force and predicted the existence of π -mesons. Many kinds of new particles have since been found, including π -mesons and muons. The field theory of π -mesons can explain the qualitative features of the meson-nucleon system. Similar theories can be formulated for other types of unstable particles that have been found in cosmic rays since 1949.

During the progress of meson theory, various types of fields were investigated, and a general theory of elementary particles was developed. Dirac (1936) proposed a general wave equation for elementary particles, and Pauli and M. Fierz (1939) proved the connection of spin and statistics (\rightarrow the end of Section D). Schwinger (1951) derived quantum-mechanical equations of motion and commutation relations from a unified variational principle. For the cases where no interaction is present, a general theory of elementary particles consistent with the requirements of relativity and of quantum theory was established as the theory of **free fields**.

B. Relativistically Covariant Classical Fields

Relativistically covariant fields $\varphi_r^\alpha(x)$ on the level of classical theory are functions of the space-time point x with either real or complex values depending on the index α (which distinguishes different fields) [1]. A finite-dimensional representation D^α of $SL(2, C)$ (\rightarrow 258 Lorentz Group) on either a real or complex vector space is assigned to each α , and the index r of $\varphi_r^\alpha(x)$ refers to components in this representation space. Each model is specified in terms of a **Lagrangian density** $\mathcal{L}(x)$ that is a function (typically a polynomial) of $\varphi_r^\alpha(x)$, $\partial_\mu \varphi_r^\alpha(x)$ (∂_μ denoting $\partial/\partial x^\mu$, $\mu=0, \dots, 3$) and their complex conjugates. \mathcal{L} is taken to be invariant under the replacement of $\varphi_r^\alpha(x)$ and $\partial_\mu \varphi_r^\alpha(x)$ by $\sum_s D^\alpha(A)_{rs} \varphi_s^\alpha(x)$ and $\sum_{s,\nu} D^\alpha(A)_{rs} \Lambda(A)_\mu^\nu \partial_\nu \varphi_s^\alpha(x)$ for all $A \in SL(2, C)$ (called the **Lorentz invariance** of \mathcal{L}) and possibly under the replacement of $\varphi_r^\alpha(x)$ and $\partial_\mu \varphi_r^\alpha(x)$ by $\sum_\beta U^{\alpha\beta} \varphi_r^\beta(x)$ and $\sum_\beta U^{\alpha\beta} \partial_\mu \varphi_r^\beta(x)$ (called an **internal symmetry**).

The fields are supposed to satisfy the following partial differential equation, called the **field equation**, and are obtained as the [†]Euler equation of the variational problem for the action integral $I = \int \mathcal{L}(x) dx$:

$$\partial \mathcal{L}(x) / \partial \varphi_r^\alpha(x) - \sum_{\mu=0}^3 (\partial / \partial x^\mu) (\partial \mathcal{L}(x) / \partial (\partial^\mu \varphi_r^\alpha(x))) = 0,$$

$$\partial \mathcal{L}(x) / \partial \bar{\varphi}_r^\alpha(x) - \sum_{\mu=0}^3 (\partial / \partial x^\mu) (\partial \mathcal{L}(x) / \partial (\partial^\mu \bar{\varphi}_r^\alpha(x))) = 0$$

(the bar indicates the complex conjugate), where the second equation is identical to the first for a real field. For complex fields, $\mathcal{L}(x)$ is chosen such that $I = \bar{I}$ (possibly except for a surface term) in order to ensure that the above two equations are complex conjugates of each other.

The invariance of I under a Lie group of (local and/or volume-preserving point) transformations of fields implies a conservation law for a certain quantity (**Noether's theorem**; \rightarrow e.g., E. L. Hill, *Rev. Mod. Phys.*, 23 (1951)). For translational invariance ($x^\mu \rightarrow x^\mu + a^\mu$), the **energy-momentum (stress) tensor**

$$T_{\nu}^{\mu}(x) \equiv -\delta_{\mu\nu} \mathcal{L}(x) + \sum_{\alpha,r} \{ \partial \mathcal{L}(x) / \partial (\partial_\mu \varphi_r^\alpha(x)) \partial_\nu \varphi_r^\alpha(x) + \partial \mathcal{L}(x) / \partial (\partial_\mu \bar{\varphi}_r^\alpha(x)) \partial_\nu \bar{\varphi}_r^\alpha(x) \}$$

(the complex conjugate terms in all equations to be suppressed for real fields) satisfies the differential conservation law $\sum_{\mu=0}^3 \partial_\mu T_{\nu}^{\mu}(x) = 0$ as a consequence of the field equations, which in turn implies the time-independence of the

following quantity provided that $T^\mu_\mu(x)$ vanishes sufficiently fast at spatial infinity:

$$H = \int T^0_0(x) d^3\mathbf{x}, \quad P_k = \int T^0_k(x) d^3\mathbf{x} \quad (x^0 = t).$$

These are the total field Hamiltonian (energy) and momentum.

In a general situation, $T^{\mu\nu} = \sum_\rho g^{\nu\rho} T^\mu_\rho$, where $g^{\mu\nu}$ is the Minkowski metric tensor with the signature 1, -1, -1, -1) is not necessarily symmetric in μ and ν . F. J. Belinfante (*Physica*, 6 (1939); 7 (1940)) has given a formula (by a change of $\mathcal{L}(x)$ via so-called 4-divergence) for obtaining a symmetrized energy-momentum tensor θ^μ_ν for which H and P_k are equal to those defined from the above T^μ_ν and for which the differential conservation law is satisfied.

Lorentz invariance implies the conservation law $\sum_\mu \partial_\mu M^{\mu\nu\lambda}(x) = 0$ for the **angular momentum density**

$$\begin{aligned} M^{\mu\nu\lambda}(x) &= x^\nu T^{\mu\lambda}(x) - x^\lambda T^{\mu\nu}(x) \\ &+ \sum_{\alpha, r, s} \{ \partial \mathcal{L} / \partial (\partial_\mu \varphi_r^\alpha(x)) S_{rs}^{\alpha\nu\lambda} \varphi_r^\alpha(x) \\ &\quad + \partial \mathcal{L} / \partial (\partial_\mu \bar{\varphi}_r^\alpha(x)) \bar{S}_{rs}^{\alpha\nu\lambda} \bar{\varphi}_r^\alpha(x) \} \\ &= x^\nu \Theta^{\mu\lambda}(x) - x^\lambda \Theta^{\mu\nu}(x) \end{aligned}$$

and the time-independence of the total angular momentum

$$M^{\nu\lambda} = \int M^{\nu\lambda} d^3\mathbf{x} \quad (x^0 = t),$$

where $S_{rs}^{\alpha\mu\nu}$ is antisymmetric in μ, ν and $D^\alpha(A)_{rs} = \delta_{rs} + \sum_{\mu, \nu} S_{rs}^{\alpha\mu\nu} \varepsilon_{\mu\nu} / 2$ for $\Lambda(A)_{\mu\nu} = g_{\mu\nu} + \varepsilon_{\mu\nu}$ with an infinitesimal $\varepsilon_{\mu\nu}$. (We have $(\Lambda(A)x)^\mu = \sum_\nu \Lambda(A)^\mu_\nu x^\nu$ and $\Lambda(A)_{\mu\nu} = \sum_\rho g_{\mu\rho} \Lambda(A)^\rho_\nu$.)

A continuous one-parameter group $U(\rho)$ of internal symmetries implies the conservation law $\sum_\mu \partial_\mu J^\mu(x) = 0$ for the **4-current density** (charge ($\mu = 0$) and current ($\mu = 1, 2, 3$) densities)

$$\begin{aligned} J^\mu(x) &= \sum_{\alpha\beta r} \{ \partial \mathcal{L} / \partial (\partial_\mu \varphi_r^\alpha(x)) \lambda^{\alpha\beta} \varphi_r^\beta(x) \\ &\quad + \partial \mathcal{L} / \partial (\partial_\mu \bar{\varphi}_r^\alpha(x)) \bar{\lambda}^{\alpha\beta} \bar{\varphi}_r^\beta(x) \} \end{aligned}$$

and the time-independence of the **charge** $\int J^0(x) d^3\mathbf{x}$ ($x^0 = t$), where $U'(0)^{\alpha\beta} = \lambda^{\alpha\beta}$. An example of $U(\rho)$ is the multiplication by $\exp iQ^z \rho$ (called a **gauge transformation of the first kind**), where Q^z is an integer with $Q^z = 0$ for real fields φ_r^z .

Examples of Lagrangian densities for non-interacting fields (called **free Lagrangian densities**) are:

real scalar field:

$$(\sum_{\mu\nu} g^{\mu\nu} \partial_\mu \varphi(x) \partial_\nu \varphi(x) - m^2 \varphi(x)^2) / 2,$$

complex scalar field:

$$(\sum_{\mu\nu} g^{\mu\nu} \partial_\mu \bar{\varphi}(x) \partial_\nu \varphi(x) - m^2 \bar{\varphi}(x) \varphi(x)),$$

Dirac field:

$$\begin{aligned} &(i/2) \sum_{rs} (\psi_r(x)^\dagger \gamma_{rs}^\mu \partial_\mu \psi_s(x) - \partial_\mu \psi_r(x)^\dagger \gamma_{rs}^\mu \psi_s(x)) \\ &\quad - m \sum_r \psi_r(x)^\dagger \psi_s(x) \\ &(\psi_r(x)^\dagger = \sum_s \bar{\psi}_s(x) \gamma_{sr}^0, \gamma^\mu \equiv (\gamma_{rs}^\mu) \text{ are } \dagger \text{Dirac's } \gamma\text{-matrices}), \end{aligned}$$

massive real vector field:

$$\begin{aligned} &-(1/4) \sum_{\kappa\lambda\mu\nu} g^{\kappa\mu} g^{\lambda\nu} (\partial_\kappa A_\lambda(x) - \partial_\lambda A_\kappa(x)) \\ &\quad \times (\partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)) + \frac{m^2}{2} \sum_{\mu\nu} g^{\mu\nu} A_\mu(x) A_\nu(x), \end{aligned}$$

where D^α is trivial for scalar fields, $[1, 0] \oplus [0, 1]$ for the Dirac field, and $[1, 1]$ for vector fields.

For interacting fields, some interaction parts are added to the sum of such free Lagrangian densities of relevant fields. Some examples are:

$P(\varphi)$ (or $g\varphi^4$) interaction:

$$P(\varphi(x)) \quad (\text{or } g\varphi(x)^4),$$

Yukawa-type interaction:

$$g \sum_{rs} \psi_r(x)^\dagger \gamma_{rs}^\mu \psi_s(x) A_\mu(x),$$

Fermi interactions:

$$\begin{aligned} &g \sum_{kk'} G_{kk'} (\sum_{rs} \psi_r^\alpha(x)^\dagger \mathcal{O}_{rs}^k \psi_s^\beta(x)) \\ &\quad \times (\sum_{r's'} \psi_{r'}^{\alpha'}(x)^\dagger \mathcal{O}_{r's'}^{k'} \psi_{s'}^{\beta'}(x)) \end{aligned}$$

($\mathcal{O}^k = 1$ (scalar, no index k), γ^μ (vector, $G_{kk'} = g_{\mu\mu'}$), $\gamma^\mu \gamma^\nu$ (tensor, $G_{kk'} = g_{\mu\mu'} g_{\nu\nu'}$), $\gamma^5 \gamma^\mu$ (pseudovector, $G_{kk'} = g_{\mu\mu'}$), $\gamma^5 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3$), γ^5 (pseudoscalar, no index k).

C. Heuristic Theory of Quantized Fields

For free fields, a quantization procedure similar to the usual quantum mechanics (\rightarrow 351 Quantum Mechanics; 377 Second Quantization) leads to the following type of canonical commutation or anticommutation relations among fields (and their first time derivatives) at time 0; for example,

real scalar field:

$$\begin{aligned} &[\varphi(0, \mathbf{x}), \dot{\varphi}(0, \mathbf{y})]_- = i\delta^3(\mathbf{x} - \mathbf{y}), \\ &[\varphi(0, \mathbf{x}), \varphi(0, \mathbf{y})]_- = [\dot{\varphi}(0, \mathbf{x}), \dot{\varphi}(0, \mathbf{y})]_- = 0, \end{aligned}$$

Dirac field:

$$\begin{aligned} &[\bar{\psi}_r(0, \mathbf{x}), \psi_s(0, \mathbf{y})]_+ = \delta_{rs} \delta^3(\mathbf{x} - \mathbf{y}), \\ &[\psi_r(0, \mathbf{x}), \psi_s(0, \mathbf{y})]_+ = [\bar{\psi}_r(0, \mathbf{x}), \bar{\psi}_s(0, \mathbf{y})]_+ = 0. \end{aligned}$$

($[A, B]_\pm = AB \pm BA$, $\delta_{rs} = 0$ for $r \neq s$ and $\delta_{rr} = 1$, $\delta^3(\mathbf{x} - \mathbf{y}) = \prod_{k=1}^3 \delta(x^k - y^k)$.) The free-field equations then lead to the following 4-dimensional commutation relations:

real scalar field:

$$[\varphi(x), \varphi(y)]_- = -i\Delta_m(x - y),$$

Dirac field:

$$[\psi_r(x), \psi_s(y)^\dagger]_+ = (\sum_\mu \gamma_{rs}^\mu \partial_\mu - im \delta_{rs}) \Delta_m(x - y),$$

massive real vector field:

$$[A_\mu(x), A_\nu(y)]_- = i(g_{\mu\nu} + m^{-2} \partial_\mu \partial_\nu) \Delta_m(x - y).$$

Here Δ_m is the \dagger invariant distribution. There is a unique representation of such relations for a

field (called the **Fock representation**) with a vector Ω (called the **free vacuum vector**) which is annihilated by $\int \varphi(x)f(x)dx$, by $\int \psi_r(x)f(x)dx$ and $\int \bar{\psi}_r(x)f(x)dx$, or by $\int A_\mu(x)f(x)dx$ whenever $\int e^{ip \cdot x} f(x)dx = 0$ for $p^0 > 0$ (here, $p \cdot x = \sum g_{\mu\nu} p^\mu x^\nu$), and which is cyclic (i.e., polynomials of (smeared-out) fields generate a dense subset of Ω). The free fields in the Fock representation satisfy the Wightman axioms (\rightarrow Section D).

In the Fock representation of canonical fields a translationally invariant vector must be a free vacuum vector, up to multiplication by a complex number. In order to construct a model of a translationally invariant interaction among canonical fields with a unique vector of minimal energy (a vector which is called the true or interacting vacuum and which must be translationally invariant if unique), one must look for some other suitable representation. Such a no-go theorem is called **Haag's theorem**.

The earliest formulation of interacting quantized fields was developed heuristically by formal manipulation in the Fock representation, described in textbooks of **quantum field theory** [2–6]. It can be mathematically justified if the so-called **cutoff** is introduced by limiting the space to a finite volume, possibly changing the space into a lattice and smoothing out fields (effectively cutting off the high-energy part of the interaction) (A. M. Jaffe, O. E. Lanford III, and A. S. Wightman, *Comm. Math. Phys.*, 15 (1969)). The full theory is then expected to be obtained by taking a limit of the true vacuum expectation values as various cutoff parameters are removed. This is the aim of constructive field theory (\rightarrow Section F), which has been achieved for some space-time models of dimension 2 and 3.

In quantum field theory, the S -matrix is given in terms of (the mass-shell restriction of the Fourier transform of) the vacuum expectation value of the time-ordered product of fields, called τ -functions (\rightarrow Section E). In the heuristic approach, it is given in terms of the following Gell-Mann–Low formula (its imaginary time version being mathematically used in constructive field theory): If the field $\varphi(x) = e^{iHx^0} \varphi_0((0, \mathbf{x})) e^{-iHx^0}$ with $H = H_0 + H_I$ (H_0 is the free Hamiltonian as the generator of the time translation for the free field $\varphi_0(x)$ and H_I is the interaction Hamiltonian, such as $P(\varphi_0)$) and $x_1^0 > \dots > x_n^0$, then

$$\begin{aligned} &(\Omega, \varphi(x_1) \dots \varphi(x_n) \Omega) \\ &= \lim_{T \rightarrow \infty} (e^{iHT} \Omega_0, \varphi(x_1) \dots \varphi(x_n) e^{-iHT} \Omega_0) \\ &\quad / (\Omega_0, e^{-2iHT} \Omega_0), \end{aligned}$$

where Ω_0 is the free vacuum. The numerator

can be written as

$$(\Omega_0, U(T, x_1^0) \varphi_0(x_1) U(x_1^0, x_2^0) \dots \varphi_0(x_n) U(x_n^0, T) \Omega_0)$$

in terms of $U(t, s) = e^{iH_0 t} e^{-i(t-s)H} e^{-isH_0}$ and the canonical field φ_0 at time 0. The covariant perturbation series due to Tomonaga, Schwinger, Feynman, and Dyson is obtained by substituting the following expansion, which is the iteration of the Duhamel formula:

$$\begin{aligned} &U(t, s) \\ &= \sum_{n=0}^{\infty} (-i)^n \int_s^t dt_1 \dots \int_s^{t_{n-1}} dt_n H_I(t_1) \dots H_I(t_n). \end{aligned}$$

Here, $H_I(t) \equiv e^{iH_0 t} H_I e^{-iH_0 t}$. Each term can be represented by connected \dagger Feynman diagrams (the denominator canceling out all disconnected graphs) and computed according to the \dagger Feynman rule, yielding \dagger Feynman integrals. Each expression so obtained (formally) may be a divergent integral (in the absence of the cutoff), in which case one tries to cancel it out by modifying the original Hamiltonian with the alteration of parameters such as mass and coupling constant (by an amount called the **renormalization constant**, which is divergent in the absence of the cutoff) or possibly by the addition of terms again involving renormalization constants. If this can be achieved in terms of a finite number of renormalization constants, the model or the Hamiltonian is called **renormalizable** [7]. If the divergent integral appears only in a finite number of graphs (counting the same subgraph of an infinite number of different diagrams as one graph), the model is called **super-renormalizable**.

D. Axiomatic Quantum Field Theory

Relativistically covariant fields $\varphi_\alpha(x)$ on a Hilbert space \mathcal{H} are called **Wightman fields** if the following four **Wightman axioms** are fulfilled [8–10]:

(1) The fields $\varphi_\alpha(x)$ are **operator-valued distributions**: For each C^∞ -function f of rapid decrease on the Minkowski space ($f \in \mathcal{S}(\mathbf{R}^4)$), $\varphi_\alpha(f)$ is an operator defined on a common domain D dense in \mathcal{H} and satisfying $\varphi_\alpha(f)D \subset D$ (the domain of $\varphi_\alpha(f)^*$ contains D), $\varphi_\alpha(f)^*|D = \varphi_{\bar{\alpha}}(\bar{f})$ for another index $\bar{\alpha}$ and $(\Psi, \varphi_\alpha(f)\Phi) = (\Psi, \varphi_\alpha(f)\Phi)$ for any Ψ , and Φ in D is linear and continuous in f relative to the topology of the Schwartz space $\mathcal{S}(\mathbf{R}^4)$.

(2) **Relativistic covariance**: There exists a continuous unitary representation $U(a, A)$ ($a \in \mathbf{R}^4$ and $A \in SL(2, \mathbf{C})$) of the universal covering group $\tilde{\mathcal{P}}_+^\uparrow$ of the \dagger restricted inhomogeneous Lorentz group \mathcal{P}_+^\uparrow on \mathcal{H} such that $U(a, A)D \subset$

D and

$$U(a, A)\varphi_\alpha(f)U(a, A)^* = \sum_\beta D(A)_{\alpha\beta}\varphi_\beta(f_{a,A}),$$

where $f_{a,A}(x) = f(\Lambda(A)^{-1}(x-a))$ and $(D(A)_{\alpha\beta})$ is a finite-dimensional representation of $SL(2, \mathbb{C})$. The index α usually consists of \dagger undotted and dotted indices, interchanged in $\bar{\alpha}$, and of other indices, interchanged among them in $\bar{\alpha}$.)

(3) **Locality:** If the supports of f and g are mutually spacelike, then

$$\varphi_\alpha(f)\varphi_\beta(g) = \varepsilon(\alpha, \beta)\varphi_\beta(g)\varphi_\alpha(f)$$

on D , where $\varepsilon(\alpha, \beta) = \pm 1$. If $\varepsilon(\alpha, \beta) = -1$ exactly when $D(A)_{\alpha\alpha} = D(A)_{\beta\beta} = -1$ for $A = -1$ (i.e., when both φ_α and φ_β are Fermi fields), they are said to satisfy the **normal commutation relations**.

(4) **Spectrum conditions:** Let $U(a, 1) = e^{ia \cdot P}$. The joint spectrum of P^μ ($\mu = 0, 1, 2, 3$) is in the forward cone $\bar{V}_+ = \{p \in \mathbb{R}^4 | p \cdot p \geq 0, p^0 \geq 0\}$, with a point spectrum of multiplicity 1 at $p = 0$. A vector Ω belonging to the point spectrum 0 of P^μ is called the true (or interacting) vacuum and is required to be in D .

Usually D is taken to be minimal, namely, D is the linear hull of Ω and $\varphi_{\alpha_1}(f_1) \dots \varphi_{\alpha_n}(f_n)\Omega$ with all possible $n, \alpha_1, \dots, \alpha_n, f_1, \dots, f_n$. By means of the nuclear theorem, it is possible to define a linear operator $\varphi_{\alpha_1 \dots \alpha_n}(f)$ for $f(x_1, \dots, x_n)$ in $\mathcal{S}(\mathbb{R}^{4n})$, linear and continuous in f on D such that it coincides with $\varphi_{\alpha_1}(f_1) \dots \varphi_{\alpha_n}(f_n)$ if $f(x) = \prod_{j=1}^n f_j(x_j)$. If such operators are introduced, then the linear hull of Ω and $\varphi_{\alpha_1 \dots \alpha_n}(f)\Omega$ is taken to be D .

Under the foregoing axiom, the vacuum expectation values of the products of field operators define tempered distributions $W_\alpha(x) = W_{\alpha_1 \dots \alpha_n}(x_1, \dots, x_n)$, called **Wightman functions**, such that

$$W_\alpha(\prod f_j) = (\Omega, \varphi_{\alpha_1}(f_1) \dots \varphi_{\alpha_n}(f_n)\Omega).$$

The notion of a connected graph in perturbation theory corresponds to the following notion of **truncated Wightman functions** W_α^T :

$$W_\alpha^T(I) = \sum_m (-1)^{m-1} (m-1)! \sum_{\{I_k\}} \prod_{k=1}^m W_\alpha(I_k),$$

$$W_\alpha(I) = \sum_m \sum_{\{I_k\}} \prod_{k=1}^m W_\alpha^T(I_k),$$

where I indicates a set of variables $x_j \in M, j \in I$, $\{I_k\}$ is a partition of I into m subsets, with the order of the x 's in each I_k remaining the same as in I , and the sum over $\{I_k\}$ extending over all possible partitions. If the spectrum of $\{P^\mu\}$ in Ω^\perp is contained in $\bar{V}_+^m = \{p \in M | p \cdot p \geq m^2, p^0 \geq m\}$ for some $m > 0$ (the mass gap), then W_α^T has an exponential clustering property at spatial infinity. For example $W_\alpha^T(x_1 + a_1, \dots, x_n + a_n)e^{m'R} \rightarrow 0$ as a distribution in x if $m' < m$, $a_j^0 = 0$ for all j and $R = \max |a_j - a_k| \rightarrow \infty$.

If the mass operator $(P \cdot P)^{1/2}$ has an isolated point spectrum at m with the eigenspace $\mathcal{H}_1(m)$, then there are sufficiently many $n, \alpha = (\alpha_1, \dots, \alpha_n)$, and $f(x_1, \dots, x_n)$ such that $\varphi_\alpha(f)\Omega \in \mathcal{H}_1(m)$, and the linear hull of such vectors are dense in $\mathcal{H}_1(m)$. In terms of the f_j satisfying $\varphi_{\alpha_j}(f_j)\Omega \in \mathcal{H}_1(m_j)$ (some of the m 's may coincide) and the solutions $g_j(x) = \int \tilde{g}_j(p) \exp i(\mathbf{p}x - \omega_j(\mathbf{p})x^0) d^3\mathbf{p}$, $\omega_j(p) = (p^2 + m_j^2)^{1/2}$, of the Klein-Gordon equation, then the following limits, called **out** and **in states**, exist:

$$\Psi^{\text{out}}(h_1 \dots h_n) = \lim_{t \rightarrow \pm\infty} Q_1(t, g_1) \dots Q_n(t, g_n)\Omega,$$

$$Q_j(t, g_j) = \int \varphi_{\alpha_j}(x_1 + (t, \mathbf{x}), \dots, x_{n_j} + (t, \mathbf{x})) \\ \times f_j(x_1 + \dots + x_{n_j}) g_j(t, \mathbf{x}) dx_1 \dots dx_{n_j} d\mathbf{x},$$

$$h_j = (2\pi)^3 \tilde{g}_j(P) \varphi_{\alpha_j}(f_j)\Omega \in \mathcal{H}_1(m_j).$$

It defines the S -matrix elements as follows:

$$S(h_1 \dots h_n; h'_1 \dots h'_n) \\ = (\Psi^{\text{out}}(h_1 \dots h_n), \Psi^{\text{in}}(h'_1 \dots h'_n)).$$

This is called **Haag-Ruelle scattering theory**, and the existence proof is based on the clustering property of W_α^T and an asymptotic estimate for the behavior of the g 's for large t and all \mathbf{x} (R. Haag, *Phys. Rev.*, 112 (1958); D. Ruelle, *Helv. Phys. Acta*, 35 (1962)). The out and in states can be interpreted in the limit of infinite future and past as the state where n particles are moving at velocities \mathbf{v}_j related to the spectrum p^μ of P^μ through the relation $p^0 = m_j/(1 - \mathbf{v}_j^2)^{1/2}$, $\mathbf{p} = m_j\mathbf{v}_j/(1 - \mathbf{v}_j^2)^{1/2}$ (with probability amplitude proportional to $h_j(p)$) (H. Araki and R. Haag, *Comm. Math. Phys.*, 4 (1967)).

Since

$$(\Psi^{\text{out}}(h_1 \dots h_n), \Psi^{\text{out}}(h'_1 \dots h'_n))$$

$$= \delta_{nn'} \sum_P \prod_{j=1}^n (h_j, h'_{P(j)}),$$

where the sum is over all permutations P , Ψ^{out} can be viewed as a unitary mapping from the \dagger Fock space \mathcal{H}^0 over $\sum_m^\oplus \mathcal{H}_1(m)$ to the closed subspace \mathcal{H}^{out} spanned by $\Psi^{\text{out}}(h_1 \dots h_n)$, $n = 0, 1, 2, \dots$ (Ω for $n = 0$, h_1 itself for $n = 1$). The free fields on \mathcal{H}^{out} and \mathcal{H}^{in} are called (out and in) **asymptotic fields**. Likewise, Ψ^{in} is a unitary mapping from \mathcal{H}^0 to \mathcal{H}^{in} . If $\mathcal{H}^{\text{in}} = \mathcal{H}^{\text{out}}$, then the matrix element S can be viewed as a matrix element of a unitary operator on \mathcal{H}^0 , called an **S -matrix**. Its unitary transform by Ψ^{out} and by Ψ^{in} coincide and define a unitary operator, sometimes called an **S -operator**, on $\mathcal{H}^{\text{in}} = \mathcal{H}^{\text{out}}$. The equality $\mathcal{H} = \mathcal{H}^{\text{in}} = \mathcal{H}^{\text{out}}$ is called the **completeness of the scattering states** or **asymptotic completeness**.

If the scattering states are complete, then the following **LSZ asymptotic condition**, due

to Lehmann, Symanzik, and Zimmermann (*Nuovo Cimento*, 1 (1955)), holds:

$$\text{weak-lim}_{t \rightarrow \pm \infty} Q_j(t, g_j) = a_{\text{out}}^*(h_j),$$

where $a^*(h)$ is the \dagger creation operator ($a^*(h) \cdot \Psi(h_1 \dots h_n) = \Psi(hh_1 \dots h_n)$ for either out or in states), $\varphi_{\alpha(j)}(f_j)\Omega \in \mathcal{H}_1(m_j)$ is not required, and $(\Omega, \varphi_{\alpha(j)}(f_j)\Omega) = 0$ is assumed instead. This leads to an explicit expression for the S -matrix elements in terms of τ -functions, which takes the following simple form if all particles have spin 0, all fields φ_α are scalar ($D(A)_{\alpha\beta} = \delta_{\alpha\beta}$), and all n_j can be taken to be 1. First define the **connected part** $S_c(h_1 \dots h_n; h'_1 \dots h'_n)$ from S by exactly the same equation as truncated Wightman functions with S to be set to 0 if $n=0$, 1 or $n'=0$, 1 except when $n=n'=1$ ($S(h, h') = (h, h')$). Then

$$\begin{aligned} S_c(h_1 \dots h_l; h_{l+1} \dots h_n) \\ = \int \left[\left(\prod_{j=1}^n (p_j \cdot p_j - m_j^2) \right) \tilde{\tau}_\alpha^T(p_1, \dots, p_l, \right. \\ \left. -p_{l+1}, \dots, -p_n) \right] \prod_{j=1}^n (-iZ_j^{-1/2} h_j(\mathbf{p}_j) d\Omega_j(\mathbf{p}_j)), \end{aligned}$$

where $p_j^0 = (\mathbf{p}_j^2 + m_j^2)^{1/2}$ (the set of such p is called the positive mass shell; there the vanishing factors $(p_j \cdot p_j - m_j^2)$ cancel the poles of $\tilde{\tau}_\alpha^T$), $d\Omega_{m_j}(\mathbf{p})$ is the Lorentz invariant measure $(2|p^0|)^{-1} d^3 \mathbf{p} = (\delta(p \cdot p - m_j^2) d^4 p)$ on the positive mass shell, $h \in \mathcal{H}_1(m)$ is represented by a measurable function $h(\mathbf{p})$ with the inner product $(h, h') = \int h(\mathbf{p}) h'(\mathbf{p}) d\Omega_m(\mathbf{p})$, Z_j is defined by $(h, \varphi_{\alpha(j)}(f_j)\Omega) = Z_j^{1/2} (h, \tilde{f}_j)$ for all $h \in \mathcal{H}_1(m_j)$ and $\tilde{f}(p) = (2\pi)^{-3/2} \int e^{ip \cdot x} f(x) dx$, and $\tilde{\tau}_\alpha^T$ is the Fourier transform of the truncated **time-ordered function** (τ -function) defined by

$$\begin{aligned} \tilde{\tau}_\alpha^T(p_1, \dots, p_n) = \int \tau_\alpha^T(x_1, \dots, x_n) \exp i \left(\sum_{j=1}^n p_j \cdot x_j \right) \\ \times \prod_{j=1}^n (2\pi)^{-3/2} d^4 x_j, \\ \tau_\alpha^T(x) = \sum_P \theta(x^0; C_P) W_{P,\alpha}^T(x) \quad (x = (x_1, \dots, x_n)), \end{aligned}$$

$$W_{P,\alpha}^T(x) = (\Omega, \varphi_{\alpha_{P(1)}}(x_{P(1)}) \dots \varphi_{\alpha_{P(n)}}(x_{P(n)}) \Omega)^T,$$

in which the sum is over all permutations P , T indicates the truncated Wightman functions, $\theta(x^0; C_P)$ is the characteristic function of the cone $x_{P(1)}^0 \geq x_{P(2)}^0 \geq \dots \geq x_{P(n)}^0$ (the time ordering) possibly smeared out by convolution with a C^∞ -function of compact support (so that it can multiply a distribution), where the formula does not actually depend on the smearing functions. (Usually fields φ_α are normalized so that $Z_j = 1$.)

The r -functions (O. Steinmann, *Helv. Phys. Acta*, 33 (1960); D. Ruelle, *Nuovo Cimento*, 19 (1961); H. Araki, *J. Math. Phys.*, 2 (1961)) are

defined by

$$\begin{aligned} r_i(x) = \sum_P \theta(x^0; C_P/C_i) W_{P,\alpha}^T(x) \\ \left(= \sum_P \theta(x^0; C_P/C_i) W_{P,\alpha}(x) \right) \end{aligned}$$

(often with an additional factor $(-i)^{n-1}$), where the Fourier-Laplace transform $\tilde{\theta}(q^0; C_P)$ of $\theta(x^0; C_P)$ is a rational function and $\theta(x^0; C_P/C_i)$ is the inverse Fourier transform of its boundary value as $\text{Im } q^0 \in C_i$ tending to 0, and the cone C_i and hence r_i is specified by a consistent choice of signs of $\sum_{j \in I} \text{Im } q_j^0$ for all nonempty proper subsets I of $(1, \dots, n)$. The Fourier transform of $r_i(x)$ is a boundary value of an analytic function common to all r_i and coincides with $\tilde{\tau}_\alpha^T$ for $p^0 \in C_i$. Making use of this relation, some analyticity properties of the S -matrix (including TCP symmetry) have been proved ([11, 12]; J. Bros, H. Epstein, and V. Glaser, *Comm. Math. Phys.*, 1 (1965); Epstein, Glaser, and A. Martin, *Comm. Math. Phys.*, 13 (1969); Epstein, *J. Math. Phys.*, 8 (1967)).

The two-point function for scalar fields has the following simple expression in terms of \dagger invariant distributions, sometimes called the **Källén-Lehmann representation** with the **Källén-Lehmann weight** ρ :

$$(\Omega, \varphi(x) \varphi(y) \Omega) = \int_0^\infty \Delta_\kappa(x-y) d\rho(\kappa^2).$$

Under the Wightman axiom (with the normal commutation relations), there exists an antiunitary operator Θ , called the **TCP operator**, which satisfies the relations

$$\Theta \Omega = \Omega, \quad \Theta U(a, A) \Theta^{-1} = U(-a, A),$$

$$\Theta^2 = U(0, -1),$$

$$\Theta \varphi_\alpha(x) \Theta^{-1} = (-1)^\eta (-i)^N \varphi_\alpha(-x),$$

where η is the number of undotted indices in α , $N=0$ if $D(A)_{\alpha\alpha} = 1$ for $A = -1$ (Bose fields), and $N=1$ if $D(A)_{\alpha\alpha} = -1$ for $A = -1$ (Fermi fields).

The choice of ± 1 for $\varepsilon(\alpha, \alpha)$ in the locality axiom cannot be opposite to the normal commutation relations for any α except for the trivial case $\varphi_\alpha(x) = 0$ ([8–10]; G. Lüders and B. Zumino, *Phys. Rev.*, 110 (1958); N. Burgoyne, *Nuovo Cimento*, 8 (1958)). This is called the **connection of spin and statistics**. For a general choice of ± 1 for $\varepsilon(\alpha, \beta)$, the Wightman axioms imply existence of a certain number of **even-oddness conservation laws** (i.e., the existence of a unitary representation u of a group $(Z_2)^l$ for some l such that $u(g)\Omega = \Omega$, $u(g)\varphi_\alpha(f)u(g)^* = \chi_\alpha(g)\varphi_\alpha(f)$ for some characters χ_α on the group) so that the **Klein transforms** $\hat{\varphi}_\alpha(x) = u(g_\alpha)\varphi_\alpha(x)$ (for some choice of g_α) of the original fields satisfy the Wightman axioms with the normal

commutation relations ([8–10]; H. Araki, *J. Math. Phys.*, 2 (1961)).

E. Theory of Local Observables

Except for the technical assumption of operator-valued distributions, physical contents of the Wightman axioms have been formulated in terms of the von Neumann algebra $\mathcal{A}(\mathcal{O})$ for bounded open sets \mathcal{O} , generated by those observables which can be measured in the space-time region \mathcal{O} , as follows:

(1) **Isotony**: If $\mathcal{O}_1 \supset \mathcal{O}_2$, then $\mathcal{A}(\mathcal{O}_1) \supset \mathcal{A}(\mathcal{O}_2)$.

(2) **Covariance**: $U(a, A)\mathcal{A}(\mathcal{O})U(a, A)^* = \mathcal{A}(\Lambda(A)\mathcal{O} + a)$.

(3) **Locality**: If $Q_j \in \mathcal{A}(\mathcal{O}_j)$ and \mathcal{O}_1 is spacelike to \mathcal{O}_2 , then $[Q_1, Q_2] = 0$. (No signal can propagate faster than the speed of light.)

In addition, the spectrum condition is as before (the stability of the vacuum) and, since we restrict our attention to the closed span of $\mathcal{A}(\mathcal{O})\Omega$ for all \mathcal{O} , Ω is assumed to be cyclic for $\bigcup_{\mathcal{O}} \mathcal{A}(\mathcal{O})$. (Then the latter is irreducible.) By treating $Q(x) = U(x, 1)QU(x, 1)^*$ for $Q \in \mathcal{A}(\mathcal{O})$ as a (noncovariant but localized) field, Haag-Ruelle scattering theory and the analyticity properties of the S -matrix described above for Wightman fields hold in exactly the same way in the theory of local observables. The notion of an algebra of local observables has been a concern of R. Haag since the late 1950s, with the consequent analogy to Wightman fields being demonstrated by Araki in his Zürich lectures of 1961–1962. Hence the above axioms are sometimes called **Haag-Araki axioms**.

With the help of the additional axiom $(\mathcal{A}(\mathcal{O}_1) \cup \mathcal{A}(\mathcal{O}_2))'' = \mathcal{A}(\mathcal{O}_1 \cup \mathcal{O}_2)$, called the **additivity**, the vacuum vector Ω is cyclic and separating for $\mathcal{A}(\mathcal{O})$ for any bounded open \mathcal{O} (**Reeh-Schlieder theorem**) and $\mathcal{A}(\mathcal{O}) = \mathcal{A}(\hat{\mathcal{O}})$, where $\hat{\mathcal{O}}$ is the double cone $\{x \mid |x^0| + |\mathbf{x}| \leq L\}$ if $\mathcal{O} = \{x \mid |x^0| + |\mathbf{x}| \leq L, |\mathbf{x}| < \varepsilon\}$ for any $\varepsilon > 0$, for example (**Borchers theorem**).

A merit of the Haag-Araki axioms is that these axioms have direct physical interpretation. In particular they always imply the commutativity at spacelike separation of supports, in contrast to the anticommutativity for Fermi fields. This then necessitates the consideration of the representations associated with some other states, such as states with an odd number of fermions and representations of the C^* -algebra \mathcal{A} of quasilocal observables (generated by all $\mathcal{A}(\mathcal{O})$), which are nonequivalent to the vacuum representation and are called **superselection sectors**. This viewpoint was introduced by Haag and D. Kastler (*J. Math. Phys.*, 5 (1964)) and the C^* -algebra

version of the Haag-Araki axioms is called the **Haag-Kastler axioms**.

If \mathcal{O}' denotes the causal complement of \mathcal{O} (i.e., the set of all points spacelike to \mathcal{O}), then the assumption $\mathcal{A}(\mathcal{O}) = \mathcal{A}(\mathcal{O}')$ is called **duality** and is proved for a certain type of region, which includes the double cone in the case of free fields. With the assumption of the duality for double cones in the vacuum sector, S. Doplicher, Haag, and J. Roberts (*Comm. Math. Phys.*, 13 and 15 (1969); 24 (1971); 35 (1974)) succeeded in the analysis of superselection sectors and clarified the connection of spin and statistics in a much more satisfactory fashion, as well as the anticommutativity of intertwining operators for superselection sectors for Fermi statistics.

F. Constructive Field Theory

An effort to make mathematical sense out of the heuristic theory of quantized fields and to produce examples of Wightman fields and the associated system of local observables has been pushed forward by J. Glimm and A. Jaffe since the mid 1960s and is known as **constructive field theory**. Since 1972, the **Euclidean methods**, already known in some sense, have become extremely powerful central tools, are collectively known as **Euclidean field theory** [13–16].

The Wightman function $W(z_1 \dots z_n)$ is analytic at the **Schwinger points** $z_j = (ix_j^0, \mathbf{x}_j)$ ($x \in \mathbf{R}^4$) if $x_j \neq x_k$ for $j \neq k$, and its value $S(x_1 \dots x_n) = W(z_1 \dots z_n)$ is called the **Schwinger function**. The axioms for Schwinger functions equivalent to Wightman axioms are known as **Osterwalder-Schrader axioms** (*Comm. Math. Phys.*, 31 (1973); 42 (1975)). The positivity axiom reflecting the positive definite metric of the Hilbert space for Wightman fields is known as **O-S positivity** (or **T-positivity** or **reflection positivity**).

Since the Schwinger function is symmetric in its variables, it can be viewed as the expectation value of the product of (commuting) random fields, called **Euclidean fields**, if an additional positivity holds. This idea was put forward by K. Symanzik in the 1960s. E. Nelson then realized that Euclidean fields for free fields have the Markov property, and he developed **Euclidean Markov field theory**. A work of Guerra in 1972 utilizing **Nelson symmetry** revealed the extreme power of this approach, and the whole of constructive field theory has been studied in Euclidean formulation with remarkable results for super-renormalizable models in space-time of dimension 2 and 3. The key point is the **Feynman-Kac-Nelson formula**, which expresses Schwinger functions

as functional integrals and reveals a mathematical connection between Euclidean field theory and classical statistical mechanics.

G. Gauge Theory

†Electrodynamics in terms of the 4-vector potential $A_\mu(x)$ is invariant under the local gauge transformations $A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \Lambda(x)$ (and the associated transformation of the charged fields) for Λ satisfying the wave equation $\square \Lambda = 0$. This leads on one hand to complication in the canonical quantization of the fields $A_\mu(x)$ and, on the other hand, to the necessity of an indefinite inner-product space as exemplified by the **Gupta-Bleuler formalism** of quantum electrodynamics. In such a formalism, the **physical Hilbert space** (with positive definite metric) is introduced by considering a specific subspace (physical subspace) of a semidefinite metric and taking the quotient by its null subspace.

The corresponding theory with a non-commutative gauge group is known as the theory of **Yang-Mills fields**. In order to restore the formal unitarity of the S -matrix in the perturbation series in terms of Feynman diagrams, L. D. Faddeev and V. N. Popov (*Phys. Lett.*, 25B (1967)) introduced fictitious particles, called **Faddeev-Popov ghosts**. T. Kugo and I. Ojima (*Prog. Theoret. Phys.*, Suppl., 66 (1979)) developed a canonical quantization scheme for the Yang-Mills field which naturally introduces additional fields corresponding to the Faddeev-Popov ghosts and specifies the physical subspace by means of the simple condition that it be the kernel of the generator of **BRS transformations**, earlier introduced by C. Becchi, A. Rouet, and R. Stora.

Gauge theory can be formulated on a lattice of finite volume as a kind of classical statistical mechanics. This is known as **lattice gauge theory**, and the important issue currently being investigated is whether or not its limit, as the lattice interval tends to 0 (the continuum limit) and the volume tends to infinity, produces a nontrivial quantum theory of gauge fields.

References

- [1] L. D. Landau and E. M. Lifshits, *The classical theory of fields*, Pergamon, 1975. (Original in Russian, 1973.)
- [2] J. D. Bjorken and S. D. Drell, *Relativistic quantum fields*, McGraw-Hill, 1965.
- [3] S. S. Schweber, *An introduction to relativistic quantum field theory*, Harper & Row, 1961.
- [4] H. Umezawa, *Quantum field theory*, North-Holland, 1956.

- [5] K. Nishijima, *Fields and particles*, Benjamin, 1969.
- [6] N. N. Bogolyubov and D. V. Shirkov, *Introduction to the theory of quantized fields*, Wiley, third edition, 1980.
- [7] K. Hepp, *Théorie de la renormalization*, Springer, 1969.
- [8] R. F. Streater and A. S. Wightman, *PCT, spin and statistics, and all that*, Benjamin, 1978.
- [9] R. Jost, *The general theory of quantized fields*, Amer. Math. Soc., 1965.
- [10] N. N. Bogolyubov, A. A. Logunov, and I. T. Todorov, *Introduction to axiomatic quantum field theory*, Benjamin, 1975. (Original in Russian, 1969.)
- [11] C. DeWitt and R. Omnes (eds.), *Dispersion relations and elementary particles*, Wiley, 1960.
- [12] M. Chretien and S. Deser (eds.), *Axiomatic field theory*, Gordon & Breach, 1966.
- [13] G. Velo and A. S. Wightman (eds.), *Constructive quantum field theory*, Springer, 1973.
- [14] B. Simon, *The $P(\varphi)_2$ Euclidean (quantum) field theory*, Princeton Univ. Press, 1974.
- [15] J. Glimm and A. Jaffe, *Quantum physics*, Springer, 1981.
- [16] B. Simon, *Functional integration and quantum physics*, Academic Press, 1979.
- [17] A. Jaffe and C. Taubes, *Vortices and monopoles*, Birkhäuser, 1980.

151 (IV.4) Finite Groups

A. The Number of Finite Groups of a Given Order

A group is called a **finite group** if its order is finite (\rightarrow 190 Groups). Since the early years of the theory of finite groups, a major problem has been to find the number of distinct isomorphism classes of groups having a given order. It is almost impossible, however, to find a general solution to the problem unless the values of n are restricted to a (small) subset of the natural numbers. Let $f(n)$ denote the number of isomorphism classes of finite groups of order n . If p is a prime number, then $f(p) = 1$ and any group of prime order is a cyclic group. If p is prime, then any group of order p^2 is an \dagger Abelian group and $f(p^2) = 2$. If p and q are distinct primes and $p > q$, then $f(pq) = 2$ or 1 according as p is congruent to 1 modulo q or not. If $p \equiv 1 \pmod{q}$, there is a non-Abelian group of order pq as well as a cyclic group of order pq . For small n , the value of $f(n)$ is as

00001
170098

follows:

n	8	12	16	18	20	24	27	28	30	32	60
$f(n)$	5	5	14	5	5	15	5	4	4	51	13

For any n , $f(n) \geq 1$. When p is prime, $f(p^m)$ is known for $m \leq 6$: $f(p^3) = 5$, $f(p^4) = 15$ if $p > 2$. For $f(p^5)$ see O. Schreier, *Abh. Math. Sem. Univ. Hamburg*, 4 (1926). For $f(2^6)$ see [11]. Set $f(p^m) = p^l$ and $l = Am^3$. Then $A \rightarrow 2/27$ as $m \rightarrow \infty$ (G. Higman, *Proc. London Math. Soc.*, 10 (1960); C. C. Sims, *Symposium on Group Theory*, Harvard, 1963).

B. Fundamental Theorems on Finite Groups

The following are some fundamental theorems useful in studying finite groups.

(1) The order of any subgroup of a finite group G divides the order of G (J. L. Lagrange). The converse is not necessarily true. If a finite group G contains a subgroup of order n for any divisor n of the order of G , then G is a \dagger solvable group. Furthermore, if G contains a unique subgroup of order n for each divisor n of the order of G , then G is a cyclic group.

Let p be a prime number. Let the order of a finite group G be $p^m m$, where m is not divisible by p . A subgroup of order p^n of G is called a **p -Sylow subgroup** (or simply a **Sylow subgroup**) of G . The importance of this concept may be seen from the next theorem.

(2) A finite group contains a p -Sylow subgroup for any prime divisor p of the order of the group. Furthermore, p -Sylow subgroups are conjugate to each other. The number of distinct p -Sylow subgroups is congruent to 1 modulo p . In general, the number of distinct p -Sylow subgroups of G that contain a given subgroup whose order is a power of p is congruent to 1 modulo p (**Sylow's theorems**).

(3) A p -group is a \dagger nilpotent group (a finite group is called a **p -group** if its order is a power of p). Thus any finite p -group G of order > 1 contains a nonidentity element in the \dagger center of G . Furthermore, any proper subgroup of G is different from its \dagger normalizer. A paper by P. Hall (*Proc. London Math. Soc.*, (2) 36 (1933)) is a classic and fundamental work on p -groups.

A group G of order 8 with 2 generators σ, τ and relations $\sigma^4 = 1$, $\tau\sigma\tau^{-1} = \sigma^{-1}$, $\sigma^2 = \tau^2$ is called the **quaternion group**. This group is isomorphic to the multiplicative group consisting of $\{\pm 1, \pm i, \pm j, \pm k\}$ in the \dagger quaternion field. A **generalized quaternion group** is a group of order 2^n with 2 generators σ, τ and relations $\sigma^{2^{n-1}} = 1$, $\tau\sigma\tau^{-1} = \sigma^{-1}$, $\tau^2 = \sigma^{2^{n-2}}$. A non-Abelian group all of whose subgroups are normal subgroups is called a **Hamilton group**. A Hamilton group is the direct product of a quaternion group, an Abelian group of odd

order, and an Abelian group of exponent 2 (i.e., $\rho^2 = 1$ for each element ρ).

Let G be a finite group, and let $G_0 = G \supset G_1 \supset \dots \supset G_r = \{1\}$ be a \dagger composition series of G . The set of isomorphism classes of the simple groups G_{i-1}/G_i , $i = 1, 2, \dots, r$, is uniquely determined (up to arrangement) by the \dagger Jordan-Hölder theorem. Thus the two most fundamental problems of finite groups are (i) the study of the simple groups and (ii) the study of a group with a given set of composition factors. The first is one of the leading problems of the theory, although it has been in a state of stagnation until rather recently (\rightarrow Section J). As to the second problem, initial works by H. Wielandt and others are under way (particularly in the direction of various generalizations of Sylow's theorems). For the class of finite solvable groups, the first problem has a rather trivial solution; only the second problem is important, and even in this case the theory seems to leave something to be desired.

C. Finite Nilpotent Groups

A finite group is nilpotent if and only if it is the \dagger direct product of its p -Sylow subgroups, where p ranges over all the prime divisors of the order. Any maximal subgroup of a nilpotent group is normal. The converse holds for finite groups; that is, a finite group is nilpotent if and only if all its maximal subgroups are normal.

D. Finite Solvable Groups

One of the most profound results on finite groups, an affirmative answer to the long-standing **Burnside conjecture**, is the **Feit-Thompson theorem** (*Pacific J. Math.*, 13 (1963)): A finite group of odd order is solvable. The index of a maximal subgroup of a finite solvable group is a power of a prime number (E. Galois). But the converse is not true. The unique simple group of order 168 has the property that all maximal subgroups are of prime power index. A finite solvable group contains a self-normalizing nilpotent subgroup (i.e., a nilpotent subgroup H such that $N_G(H) = H$), and any two such subgroups are conjugate (R. W. Carter, *Math. Z.*, 75 (1960); cf. W. Gaschütz, *Math. Z.*, 80 (1963), for a generalization). Such a subgroup is called a **Carter subgroup** and is an analog of a Cartan subalgebra of a Lie algebra. But unlike Cartan subalgebras, most simple groups do not contain any self-normalizing nilpotent subgroups. A finite solvable group of order mn (m, n are relatively prime) contains a subgroup of order m ; two subgroups of order m are conjugate; if l

is a divisor of m , then any subgroup of order l is contained in a subgroup of order m (P. Hall). The converse of the first part of this theorem is also true: A finite group is solvable if it contains a subgroup of order m for any decomposition of the order in the form mn , $(m, n) = 1$ (P. Hall's solvability criterion). This generalizes the famous **Burnside theorem** asserting the solvability of a group of order $p^a q^b$, where both p and q are prime numbers. If the sequence of the quotient groups of a principal series of a finite group G consists of cyclic groups, then the group G is called **supersolvable**. A finite group is supersolvable if and only if the index of any maximal subgroup is a prime number (B. Huppert, *Math. Z.*, 60 (1954)). If p is the largest prime divisor of the order of a finite supersolvable group G , then a p -Sylow subgroup of G is a normal subgroup.

E. Hall Subgroups

A subgroup is called a **Hall subgroup** if its order is relatively prime to its index (see the theorems of P. Hall on finite solvable groups). There is no general theorem known on the existence of a Hall subgroup. If a finite group G has a normal Hall subgroup N , then G contains a Hall subgroup H that is a complement of N in the sense that $G = NH$ and $N \cap H = 1$; furthermore, any two complements are conjugate (**Schur-Zassenhaus theorem**). The analog of Hall's theorem on finite solvable groups fails for nonsolvable groups. But if a finite group, solvable or not, contains a nilpotent Hall subgroup H of order n , then any subgroup of an order dividing n is conjugate to a subgroup of H (Wielandt, *Math. Z.*, 60 (1954); cf. P. Hall, *Proc. London Math. Soc.*, 4 (1954), for a generalization). There are some generalizations of these results for maximal π -subgroups which may not be Hall subgroups.

F. π -Solvable Groups

Let π be a set of prime numbers. Denote the set of prime numbers not in π by π' . A finite group is called a **π -group** if all the prime divisors of the order belong to π . A finite group is called **π -solvable** if any composition factor is either a π' -group or a solvable π -group. If $\pi = \{p\}$ consists of a single prime number p , we use terms such as p -group or p -solvable (instead of $\{p\}$ -solvable). Let G be a π -solvable group. A series of subgroups $P_0 = 1 \subset N_0 \subsetneq P_1 \subsetneq N_1 \subsetneq \dots \subsetneq P_l \subset N_l = G$ defined by the properties that P_i/N_{i-1} is the maximal normal π -subgroup of G/N_{i-1} and N_i/P_i is the maximal normal π' -subgroup of G/P_i is called the **π -series** of G , and the integer l is called the π -

length of G . A solvable group is π -solvable for any set π of prime numbers. A π -solvable group contains a Hall subgroup which is a π -group and also a Hall subgroup which is a π' -group; an analog of Hall's theorem on finite solvable groups holds. Hall and Higman (*Proc. London Math. Soc.*, 7 (1956)) discovered deep relations between the p -length of a p -solvable group and invariants of its p -Sylow subgroup. For example, the p -length is 1 if a p -Sylow subgroup is Abelian.

G. Permutation Groups

The set of all permutations on a set Ω of n elements forms a group of order $n!$ whose structure depends only on n . This group is called the **symmetric group of degree n** , denoted by S_n . Any subgroup of S_n is a **permutation group of degree n** . When it is necessary to mention the set Ω on which permutations operate, S_n may be denoted as $S(\Omega)$, and a subgroup of $S(\Omega)$ is called a permutation group on Ω . The set of n elements on which S_n operates is usually assumed to be $\{1, 2, \dots, n\}$, and an element σ of S_n is written as

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ 1' & 2' & \dots & n' \end{pmatrix},$$

$$\text{for example, } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 4 & 6 \end{pmatrix},$$

where i' is the image of i by σ : $i' = \sigma(i)$. The element σ may be written as $(\dots)(abc\dots z)(\dots)$, which means that σ cyclically maps a into b , b into c , and so on, and finally z back into a . In the example, $\sigma = (1\ 2\ 3)(4\ 5)(6)$. It is customary to omit the cycle with only one letter in it, such as (6) in the example. With this convention, $\sigma = (1\ 2\ 3)(4\ 5)$ may be an element of S_n for any $n \geq 5$, leaving all the letters $i \geq 6$ invariant. A **cycle** of length l is an element σ of S_n which moves l letters a_1, \dots, a_l cyclically and fixes all the rest; i.e., $\sigma = (a_1, \dots, a_l)$. Then an expression such as $\sigma = (1\ 2\ 3)(4\ 5)$ is the same as the product of two cycles $(1\ 2\ 3)$ and $(4\ 5)$. In general, any permutation can be expressed as the product of mutually disjoint cycles (two cycles (a_1, \dots, a_l) and (b_1, \dots, b_m) are said to be disjoint if $a_i \neq b_j$ for all i and j). Furthermore, the expression of the permutation as the product of mutually disjoint cycles is unique up to the order in which these cycles are written. A cycle of length 2 is called a **transposition**. Any permutation may be written as a product of transpositions. This expression is not unique, but the parity of the number of transpositions in the expression is determined by the permutation. A permutation is called **even** if it is the product of an even number of transpositions

and **odd** otherwise. The symmetric group S_n contains the same number of even and odd permutations.

The totality of even permutations forms a normal subgroup of order $(n!)/2$ ($n \geq 2$), called the **alternating group of degree n** and usually denoted by A_n . An even permutation is the product of cycles of length 3. The alternating group A_5 of degree 5 is the nonsolvable group of minimal order. This fact was known to Galois. If $n \neq 4$, then the alternating group A_n is a simple group and a unique proper normal subgroup of S_n . If $n = 4$, A_4 contains a normal noncyclic subgroup V of order 4. In this case A_4 and V are the only proper normal subgroups of S_4 . A noncyclic group of order 4 is called a (**Klein**) **four-group**. If $n \geq 5$, the symmetric group S_n is not a solvable group. This is the group-theoretic ground for the famous theorem, proved by Ruffini, Abel, and Galois, which asserts the impossibility of an algebraic solution of a general algebraic equation of degree more than four. If $n \leq 4$, S_n is solvable. The group S_4 has a composition series with composition factors of orders 2, 3, 2, 2, and S_4 is realized as the group of motions in 3-dimensional space which preserve an octahedron. Hence S_4 is called the **octahedral group**. Similarly, $A_4(A_5)$ is realized as the group of motions in space which preserve a tetrahedron (icosahedron); thus A_4 is called the **tetrahedral group** and A_5 the **icosahedral group**.

These groups have been extensively studied in view of their geometric aspect. The group of motions of a plane which preserve a regular polygon is called a **dihedral group**. If a regular polygon has n sides, then the group has order $2n$. Sometimes a Klein four-group is included in the class of dihedral groups (for $n = 2$). The dihedral groups, octahedral group, etc., are called **regular polyhedral groups**. A finite subgroup of the group of motions in 3-dimensional space is either cyclic or one of the regular polyhedral groups. A dihedral group is generated by two elements of order 2. Conversely, a finite group generated by two elements of order 2 is a dihedral group. This simple fact has surprisingly many consequences in the theory of finite groups of even order [15, ch. 9]. A dihedral group of order $2n$ contains a cyclic normal subgroup of order n , and hence is solvable.

If $n \neq 6$, every automorphism of S_n is inner. The order of the group of automorphisms of S_6 is twice the order of S_6 . The index $(S_n : H)$ of a subgroup H of S_n is at least equal to n unless $H = A_n$. If $(S_n : H) = n$, then H is isomorphic to S_{n-1} . If $n \neq 6$, S_n contains a unique conjugate class of subgroups of index n . But S_6 contains two such classes, which are exchanged by an automorphism of S_6 (\rightarrow Section I).

H. Transitive Permutation Groups

A permutation group G on a set Ω is called a **transitive permutation group** if for any pair (a, b) of elements of Ω , there exists a permutation of G which sends a into b . Otherwise G is said to be **intransitive**. Let G be a transitive permutation group on a set Ω , and let a be an element of Ω . The totality of elements of G which leave a invariant forms a subgroup of G called the **stabilizer of a** (in G). The index of the stabilizer is equal to the number of elements of Ω , the degree of G . Thus the degree of a transitive permutation group G divides the order of G (a fundamental theorem).

The concept of orbits is important. Let G be a permutation group on a set Ω . A subset Γ of Ω is called an **orbit** of G if it is G -invariant and G acts transitively on Γ . In other words, a subset Γ of Ω is an orbit of G if the following two conditions are satisfied: (i) If $a \in \Gamma$ and $g \in G$, the image $g(a)$ also lies in Γ ; and (ii) if a and b are two elements of Γ , there exists an element x of G such that $b = x(a)$. Thus each element x of G induces a permutation $\varphi_x(x)$ on Γ . The set of all these permutations $\varphi_x(x \in G)$ forms a permutation group on Γ , which may be denoted by $\varphi_\Gamma(G)$. Then $\varphi_\Gamma(G)$ is transitive on Γ , and φ_Γ is a homomorphism of G onto $\varphi_\Gamma(G)$. Thus the number of elements in an orbit Γ is a divisor of the order of G . It is clear that the set Ω on which G acts is the union of mutually disjoint orbits $\Gamma_1, \dots, \Gamma_r$ of G . This implies that the degree of G is the sum of the numbers of elements in the orbits Γ_i . The resulting equation often contains non-trivial relations. If φ_i denotes the homomorphism φ_{Γ_i} defined before, then G is isomorphic to a subgroup of the direct product of the groups $\varphi_i(G)$, $i = 1, 2, \dots, r$.

A transitive permutation group is called **regular** if the stabilizer of any letter is the identity subgroup $\{1\}$. A transitive permutation group is regular if and only if its order equals its degree. Any group can be realized as a regular permutation group (**Cayley's theorem**). A transitive permutation group which is Abelian is always regular.

Let G be a transitive permutation group on a set Ω . If the stabilizer of an element a of Ω is a maximal subgroup, G is called **primitive**, and otherwise **imprimitive**. A normal subgroup, which is $\neq \{1\}$, of a primitive permutation group is transitive. An imprimitive permutation group induces a decomposition of the set Ω into the union of mutually disjoint subsets $\Delta_1, \dots, \Delta_s$ ($s > 1$) such that each Δ_i contains at least two elements, and if $x \in G$ maps an element a of Δ_i onto an element b of Δ_j , then x maps every element of Δ_i into Δ_j ; $x(\Delta_i) = \Delta_j$. The set $\{\Delta_1, \dots, \Delta_s\}$ is called a system of im-

primitivity. A subset Δ of Ω is called a **block** if $x(\Delta) \cap \Delta$ equals Δ or the empty set for all x in G . A block is called nontrivial if $\Delta \neq \Omega$ and Δ contains at least two elements. Each member of a system of imprimitivity is a nontrivial block. A transitive permutation group is primitive if and only if there is no nontrivial block.

A permutation group G on a set Ω is called **k -transitive** (or **k -ply transitive**, where k is a natural number) if for two arbitrary k -tuples (a_1, \dots, a_k) and (b_1, \dots, b_k) of distinct elements of Ω , there is an element of G which maps a_i into b_i for all $i = 1, 2, \dots, k$. If $k \geq 2$, G is called **multiply transitive**. A doubly transitive permutation group is always primitive. The symmetric group S_n of degree n is n -transitive, while the alternating group A_n is $(n-2)$ -transitive for $n \geq 3$. Conversely, an $(n-2)$ -transitive permutation group on $\{1, 2, \dots, n\}$ is either S_n or A_n . For multiply transitive permutation groups which are simple, see the list in Section I. If $k \geq 6$, no k -transitive groups are known at present except S_n and A_n . If the Schreier conjecture (\rightarrow Section I) is true, then there are no 6-transitive permutation groups except S_n and A_n (Wielandt, *Math. Z.*, 74 (1960); H. Nagao, *Nagoya J. Math.*, 27 (1964); O'Nan, *Amer. Math. Soc. Notices*, 20 (1973)).

Two 5-transitive permutation groups other than S_n and A_n are known: the groups M_{12} and M_{24} of degrees 12 and 24, respectively, discovered by E. L. Mathieu in 1864 and 1871. The stabilizer of a letter in $M_{12}(M_{24})$ is a 4-transitive permutation group of degree 11 (23), denoted by $M_{11}(M_{23})$. No 4-transitive permutation groups other than S_n , A_n , M_i ($i = 11, 12, 23$, and 24) are known. The groups M_{12} , M_{11} , M_{24} , M_{23} and the stabilizer M_{22} of a letter in M_{23} are called **Mathieu groups**. They are simple groups which have quite exceptional properties. For Mathieu groups, see E. Witt, *Abh. Math. Sem. Univ. Hamburg*, 12 (1938). A k -transitive permutation group G on Ω of degree n and order $n(n-1)\dots(n-k+1)$ has the property that no nonidentity element of G leaves k distinct letters of Ω invariant. If $k \geq 4$, such a group is one of the following: S_k , A_{k+2} , M_{12} , and M_{11} (C. Jordan). For $k = 2$ and 3, see H. Zassenhaus, *Abh. Math. Sem. Univ. Hamburg*, 11 (1936).

A multiply transitive permutation group G contains a normal subgroup S such that S is a non-Abelian simple group and G is isomorphic to a subgroup of the group $\text{Aut } S$ of the automorphisms of S , except when the degree n of G is a power of a prime number and G contains a regular normal subgroup of order n which is an elementary Abelian group (W. S. Burnside). Furthermore, in these exceptional cases,

G is at most 2-transitive if n is odd, while G is at most 3-transitive if n is even but more than 4. The symmetric group S_4 of degree 4 is the only 4-transitive group that contains a proper solvable normal subgroup.

A transitive extension of a permutation group H on Ω is defined as follows. Let ∞ be a new element not contained in Ω . A **transitive extension** G of H is a transitive permutation group on the set $\{\Omega, \infty\}$ in which the stabilizer of ∞ is the given permutation group H on Ω . Transitive extensions do not exist for some H . Suppose that a permutation group H admits a transitive extension G that is primitive. If H is simple, then G is also simple unless the degree of G is a power of a prime number. Constructing transitive extensions has been an effective method for constructing sporadic simple groups.

Permutation groups of prime degree have been studied extensively since the last century, partly because of their connection with algebraic equations of prime degree. Let p be a prime number. A transitive permutation group of degree p is either multiply transitive and nonsolvable or has a normal subgroup of order p with factor group isomorphic to a cyclic group of order dividing $p-1$ (Burnside). Choose two cycles x and y of length p in S_p . If y is not a power of x , then the subgroup $\langle x, y \rangle$ generated by x and y is a multiply transitive permutation group which is simple. The structure of $\langle x, y \rangle$ is not known despite its simple definition. More attention has been paid to groups of degree p , where p is a prime number such that $(p-1)/2 = q$ is another prime number. The problem is to decide if such nonsolvable groups contain the alternating group A_p . The Mathieu groups M_{11} and M_{23} are the only known exceptions for $p > 7$. The search for additional exceptions has been aided by the development of high-speed computers. It is known that there is no exceptional group of degree $p = 2q + 1$ for $23 < p \leq 4079$ (P. J. Nikolai and E. T. Parker, *Math. Tables Aids Comput.*, 12 (1958); see N. Ito, *Bull. Amer. Math. Soc.*, 69 (1963), for further results in this direction).

A **Frobenius group** is a nonregular transitive permutation group in which the identity is the only element leaving more than one letter invariant. A Frobenius group of degree n contains exactly $n-1$ elements which displace all the letters. These $n-1$ elements together with the identity form a regular normal subgroup of order n . This is a theorem of Frobenius; all the existing proofs depend on the theory of characters. The regular normal subgroup of a Frobenius group is nilpotent (J. G. Thompson, *Proc. Nat. Acad. Sci. US*, 45

(1959); [15, ch. 10; 16, ch. 3]). A **Zassenhaus group** is a transitive extension of a Frobenius group.

I. Finite Simple Groups

All simple groups of finite order were completely classified in February 1982 (\rightarrow Section J; [23]). These are divided into the following four classes: (1) cyclic groups of prime order, (2) alternating groups of degree ≥ 5 , (3) simple groups of Lie type, and (4) other simple groups.

The subclass (1) consists of cyclic groups of prime order p for any prime number p . Abelian simple groups belong to this subclass. The subclass (4) consists of twenty-six sporadic simple groups including five Mathieu groups. All sporadic groups, other than the five Mathieu groups, are of recent discovery.

Simple groups of Lie type are analogs of simple Lie groups, and include the classical groups as well as the exceptional groups and the groups of twisted type.

Classical groups are divided into four types: t linear, t unitary, t symplectic, and t orthogonal (\rightarrow 60 Classical Groups). Let $q = p^r$ be a power of a prime number p . Consider a vector space V of dimension $n \geq 2$ over the field \mathbb{F}_q of q elements, except in the unitary case where V is a vector space of dimension $n \geq 2$ over the field \mathbb{F}_{q^2} of q^2 elements. Let f be a nondegenerate form on V which is t Hermitian in the unitary case (with respect to the automorphism of order 2 of \mathbb{F}_{q^2} over \mathbb{F}_q), t skew symmetric bilinear in the symplectic case, and t quadratic in the orthogonal case. In the orthogonal case, the dimension of V is assumed ≥ 3 . Consider the group of all linear transformations of V (linear case) of determinant 1, or the group of all linear transformations of determinant 1 which leave the form f invariant (in other cases). In the orthogonal case, take the commutator subgroup. With each of these groups, the factor group of it by its center is a simple group with a few exceptions.

There are several notations to denote these groups. E. Artin's notation for simple groups, which is reasonably descriptive and simple, follows the name of the simple group: n and q are as described in the preceding paragraph, q is the order of the simple group, and (a, b) denotes the greatest common divisor of two natural numbers a and b .

Linear simple group, $L_n(q)$:

$$g = q^{n(n-1)/2} \prod_{i=2}^n (q^i - 1)/d, \quad d = (n, q-1).$$

Unitary group, $U_n(q)$:

$$g = q^{n(n-1)/2} \prod_{i=2}^n (q^i - (-1)^i)/d, \quad d = (n, q+1).$$

The structure of unitary groups does not depend on the form.

Symplectic group, $S_n(q)$, $n = 2m$:

$$g = q^{m^2} \prod_{i=1}^m (q^{2i} - 1)/d, \quad d = (2, q-1).$$

In the symplectic case, the dimension n of the space V must be even, so $n = 2m$, and the structure does not depend on the form.

Orthogonal group in odd dimension $n = 2m+1$, $O_{2m+1}(q)$:

$$g = q^{m^2} \prod_{i=1}^m (q^{2i} - 1)/d, \quad d = (2, q-1).$$

The structure does not depend on the form in odd dimension.

Orthogonal groups in even dimension $n = 2m$: There are two inequivalent forms, one with index m (which is maximal) and the other with index $m-1$. The two orthogonal groups are denoted by $O_{2m}(\varepsilon, q)$, $\varepsilon = \pm 1$, where $\varepsilon = 1$ if the form is of maximal index and -1 otherwise. Then

$$g = q^{m(m-1)}(q^m - \varepsilon) \prod_{i=1}^{m-1} (q^{2i} - 1)/d, \\ d = (4, q^m - \varepsilon).$$

The value of ε is determined by the form f : $\varepsilon = 1$ if f is equivalent to $\sum_{i=1}^m x_{2i-1} x_{2i}$, and $\varepsilon = -1$ if $f \sim x_1^2 + \gamma x_1 x_2 + x_2^2 + \sum_{i=2}^m x_{2i-1} x_{2i}$, where the polynomial $t^2 + \gamma t + 1$ is irreducible over \mathbb{F}_q .

There are other ways to denote these groups. Let $X = X(*, *)$ be a group of nonsingular linear transformations of a vector space V . Two asterisks indicate two invariants, such as the dimension of V and the number of elements in the ground field. The notation SX stands for the subgroup of X consisting of linear transformations with determinant 1, and the notation PX stands for the factor group of the linear group X by its center. Thus PX is a subgroup of the group of all projectivities of the projective space formed by the linear subspaces of V . The following list is self-explanatory, except the last term in each row, which is the notation of L. E. Dickson [1]:

$$L_n(q) = PSL(n, q) = LF(n, q)$$

$$U_n(q) = PSU(n, q) = HO(n, q^2)$$

$$S_{2n}(q) = PSp(n, q) = A(2n, q).$$

(LF : linear fractional group; HO : hyperorthogonal group; A : Abelian linear group.) If f is a nondegenerate quadratic form, then the sub-

group of $GL(n, q)$ consisting of all the elements leaving the form f invariant may be denoted by $O(n, q, f)$. Let $\Omega(n, q, f)$ denote the commutator subgroup of $O(n, q, f)$. Set $\varepsilon = 1$ if f is of maximal index, and $\varepsilon = -1$ otherwise. Then

$$O_n(\varepsilon, q) = P\Omega(n, q, f).$$

Dickson's notation for orthogonal groups is complicated and seldom used.

Finite simple groups corresponding to Lie groups of some exceptional type were studied by Dickson early in this century, but C. Chevalley (*Tôhoku Math. J.*, (2) 7 (1955)) proved the existence, simplicity, and other properties of groups of any (exceptional) type over any field by a unified method. Simple Lie algebras over the field \mathbb{C} of complex numbers are completely classified, and according to the classification theory they are in one-to-one correspondence with the \dagger Dynkin diagrams. Let L be a simple Lie algebra (over \mathbb{C}) corresponding to the \dagger Dynkin diagram of type X (\rightarrow 248 Lie Algebras). Let $L = L_0 + \sum L_\alpha$ be a Cartan decomposition of L , where α ranges over the \dagger root system Δ of L . It is possible to choose a basis B of L with the following properties (\dagger Chevalley's canonical basis): B consists of $e_\alpha \in L_\alpha$ ($\alpha \in \Delta$) and a basis of L_0 ; the structure constants of L with respect to B are all rational integers; the automorphism $x_\alpha(\xi)$ in the \dagger adjoint group defined by

$$x_\alpha(\xi) = \exp(\xi \operatorname{ad} e_\alpha) \quad (\xi \in \mathbb{C})$$

maps each element of B into a linear combination of elements of B with coefficients which are polynomials in ξ with integer coefficients. Thus the matrix $A_\alpha(\xi)$ representing the transformation $x_\alpha(\xi)$ with respect to B has coefficients which are polynomials in ξ with integer coefficients.

The elements of B span a Lie algebra L_Z over the ring \mathbb{Z} of integers. Let F be a field and form $L_F = F \otimes_{\mathbb{Z}} L_Z$. Then L_F is a Lie algebra over F , and the set B may be identified with a basis of L_F over F . Let t be any element of F , $A_\alpha(t)$ be the matrix obtained from $A_\alpha(\xi)$ by replacing the complex variable ξ by the element t , and finally $x_\alpha(t)$ be the linear transformation of L_F represented by the matrix $A_\alpha(t)$ with respect to B . The group generated by the $x_\alpha(t)$ for each root α and each element t of F is called the **Chevalley group** of type X over F . The commutator subgroups of the Chevalley groups are simple, with a few exceptions which will be stated after the complete list of simple groups of Lie type. Suppose that $X = A_n, D_n$, or E_6 (\rightarrow 248 Lie Algebras S). Then the Dynkin diagram of type X has a nontrivial symmetry; let it be $\alpha \rightarrow \beta$. Suppose that the field F has an automorphism σ of the same order as the order of the symmetry of the

Dynkin diagram. Let θ be the automorphism of the Chevalley group which sends $x_\alpha(t)$ to $x_\beta(t^\sigma)$. Let U (resp. V) be the subgroup of the Chevalley group generated by $x_\alpha(t)$ with $\alpha > 0$, $t \in F$ ($x_\beta(t)$, $\beta < 0$), and let U^1 (V^1) be the subgroup consisting of all the elements of U (V) which are left invariant by θ . The group generated by U^1 and V^1 is called the **group of twisted type**. If the order of σ is i , this group is said to be of twisted type iX . In all but one case, the group of twisted type is simple (see R. Steinberg, *Pacific J. Math.*, 9 (1959)). The value i is 2 except when $X = D_4$. Since D_4 admits symmetries of orders 2 and 3, there are two twisted types. If $X = B_2, G_2$, or F_4 , then the diagram has a symmetry. If the characteristic p of the ground field F is 2, 3, or 2 according as $X = B_2, G_2$, or F_4 and the field F has an automorphism σ such that $(t^\sigma)^\sigma = t^p$ for any $t \in F$, then a procedure similar to the one described before is applicable, and the group of twisted type X' is obtained (R. H. Ree, *Amer. J. Math.*, 83 (1961)). The group of twisted type is simple if the field F has more than three elements.

The following list contains all the simple groups of Lie type. For each classical group, we list the type followed by identification:

$$A_n = L_{n+1}(q) \quad (n \geq 1)$$

$${}^2A_n = U_{n+1}(q) \quad (n \geq 1)$$

$$B_n = O_{2n+1}(q) \quad (n > 1)$$

$$C_n = S_{2n}(q) \quad (n > 2)$$

$$D_n = O_{2n}(1, q) \quad (n > 3)$$

$${}^2D_n = O_{2n}(-1, q) \quad (n > 3)$$

For other groups the type of the group is followed by the customary name or notation, if any, and the order q :

$$B'_2 \quad \textbf{Suzuki group, Sz}(q), q = 2^{2n+1}$$

$$g = q^2(q-1)(q^2+1)$$

$${}^3D_4 \quad g = q^{12}(q^8+q^4+1)(q^6-1)(q^2-1)$$

$$G_2 \quad g = q^6(q^6-1)(q^2-1)$$

$$G'_2 \quad \textbf{Ree group, Re}(q), q = 3^{2n+1}$$

$$g = q^3(q^3+1)(q-1)$$

$$F_4 \quad g = q^{24}(q^{12}-1)(q^8-1)(q^6-1) \\ \times (q^2-1)$$

$$F'_4 \quad q = 2^{2n+1} \quad g = q^{12}(q^6+1)(q^4-1) \\ \times (q^3+1)(q-1)$$

$$E_6 \quad dg = q^{36}(q^{12}-1)(q^9-1)(q^8-1) \\ \times (q^6-1)(q^5-1)(q^2-1)$$

$$d = (3, q-1)$$

$$\begin{aligned}
{}^2E_6 \quad dg &= q^{36}(q^{12}-1)(q^9+1)(q^8-1) \\
&\quad \times (q^6-1)(q^5+1)(q^2-1) \\
d &= (3, q+1) \\
E_7 \quad dg &= q^{63}(q^{18}-1)(q^{14}-1)(q^{12}-1) \\
&\quad \times (q^{10}-1)(q^8-1)(q^6-1)(q^2-1) \\
d &= (2, q-1) \\
E_8 \quad g &= q^{120}(q^{30}-1)(q^{24}-1)(q^{20}-1) \\
&\quad \times (q^{18}-1)(q^{14}-1)(q^{12}-1) \\
&\quad \times (q^8-1)(q^2-1).
\end{aligned}$$

B'_2 : M. Suzuki, *Proc. Nat. Acad. Sci. US*, 46 (1960); G'_2 and F'_4 : Ree, *Amer. J. Math.*, 83 (1961); G_2 : Dickson, *Trans. Amer. Math. Soc.*, 2 (1901), *Math. Ann.*, 60 (1905); other Chevalley groups: Chevalley, *Tôhoku Math. J.*, (2) 7 (1955); twisted types: Steinberg, *Pacific J. Math.*, 9 (1959), J. Tits, *Séminaire Bourbaki* (1958), *Publ. Math. Inst. HES* (1959), D. Hertzig, *Amer. J. Math.*, 83 (1961), *Proc. Amer. Math. Soc.*, 12 (1961).

Nonsimple cases: $L_2(2)$, $L_2(3)$, $U_3(2)$, and $Sz(2)$ are solvable groups of orders 6, 12, 72, and 20, respectively. The groups $O_3(2)$, $G_2(2)$, $G'_2(3)$, and $F'_4(2)$ contain normal subgroups of indices 2, 2, 3, and 2, respectively. These normal subgroups are simple and identified as $L_2(9)$, $U_3(3)$, $L_2(8)$ in the first three cases. The normal subgroup of $F'_4(2)$ is not in the list of simple groups of Lie type and is quite exceptional (**Tits's simple group**, *Ann. Math.*, (2) 80 (1964)).

Isomorphisms between various simple groups: $L_2(q) = U_2(q) = S_2(q) = O_3(q)$; $O_5(q) = S_4(q)$; $O_4(1, q) = L_2(q) \times L_2(q)$; $O_4(-1, q) = L_2(q^2)$; $O_6(1, q) = L_4(q)$; $O_6(-1, q) = U_4(q)$; $O_{2n+1}(q) = S_{2n}(q)$ if q is a power of 2; $L_2(2) = S_3$; $L_2(3) = A_4$; $L_2(4) = L_2(5) = A_5$; $L_2(7) = L_3(2)$; $L_2(9) = A_6$; $L_4(2) = A_8$; $U_4(2) = S_4(3)$. If q is odd and $2n \geq 6$, then $S_{2n}(q)$ and $O_{2n+1}(q)$ have the same order but are not isomorphic. $L_3(4)$ and $L_4(2)$ have the same order but are not isomorphic. There is no other isomorphism or coincidence of orders among the known simple groups (Artin, *Comm. Pure Appl. Math.*, 8 (1955)).

The groups of the automorphisms of simple groups belonging to subclasses (1), (2), and (3) are known. For the simple groups of Lie type, see Steinberg, *Canad. J. Math.*, 10 (1960).

The following list of twenty-six groups consists of all the simple groups that belong to class (4):

Five Mathieu groups whose orders are

$$M_{11}: g = 7,920 = 2^4 \cdot 3^2 \cdot 5 \cdot 11$$

$$M_{12}: g = 95,040 = 2^6 \cdot 3^3 \cdot 5 \cdot 11$$

$$M_{22}: g = 443,520 = 2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11$$

$$M_{23}: g = 10,200,960 = 2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 23$$

$$M_{24}: g = 244,823,040 = 2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23.$$

The other twenty-one groups have been discovered since 1964. Each group is identified by the symbol $(x)_i$, indicating that it is the i th group discovered in the year 19x. The list continues with the name or names of discoverers, the order of the group, and a brief description.

$(64)_1$: Z. Janko, $g = 175,560 = 2^3 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19$. A subgroup of the Chevalley group $G_2(11)$. See *J. Algebra*, 3 (1966).

$(67)_1$: M. Hall and Z. Janko, $g = 604,800 = 2^7 \cdot 3^3 \cdot 5^2 \cdot 7$, a transitive extension of $U_3(3)$ of degree 100.

$(67)_2$: D. G. Higman and Sims, $g = 44,352,000 = 2^9 \cdot 3^2 \cdot 5^3 \cdot 7 \cdot 11$, a transitive extension of M_{22} of degree 100. It is a normal subgroup of index 2 in the group of automorphisms of a certain graph with 100 vertices.

$(67)_3$: Suzuki, $g = 448,345,497,600 = 2^{13} \cdot 3^7 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13$, a transitive extension of $G_2(4)$; defined from the automorphism group of a graph of 1782 vertices.

$(67)_4$: J. McLaughlin, $g = 898,128,000 = 2^7 \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11$, a transitive extension of $U_4(3)$; defined from a graph of 275 vertices.

$(68)_1$: G. Higman, Z. Janko, and J. McKay, $g = 50,232,960 = 2^7 \cdot 3^5 \cdot 5 \cdot 17 \cdot 19$, a transitive extension of the group which is obtained from $L_2(16)$ by adjoining the field automorphism of order 2. The existence was verified by using a computer.

$(68)_2$, $(68)_3$, $(68)_4$: J. H. Conway,

$$g = 2^{21} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23$$

$$= 4,157,776,806,543,360,000,$$

$$g = 2^{18} \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23,$$

$$g = 2^{10} \cdot 3^7 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23.$$

The big group is obtained from the automorphism group of a lattice in 24-dimensional space, and the two smaller ones are subgroups of it. The lattice was defined by J. Leech in connection with a problem of close packing of spheres in 24 dimensions (*Canad. J. Math.*, 19 (1967)).

$(68)_5$: B. Fischer, $g = 2^{17} \cdot 3^9 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 = 70,321,751,654,400$, a transitive extension of $U_6(2)$ derived by means of a certain graph.

$(69)_1$: D. Held and others, $g = 2^{10} \cdot 3^3 \cdot 5^2 \cdot 7^3 \cdot 17 = 4,030,387,200$.

$(69)_2$: B. Fischer, $g = 2^{18} \cdot 3^{13} \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 23 = 4,089,470,473,293,004,800$.

$(69)_3$: B. Fischer, $g = 2^{21} \cdot 3^{16} \cdot 5^2 \cdot 7^3 \cdot 11 \cdot 13 \cdot 17 \cdot 23 \cdot 29 = 1,255,205,709,190,661,721,292,800$.

$(71)_1$: R. N. Lyons and C. C. Sims, $g = 2^8 \cdot 3^7 \cdot 5^6 \cdot 7 \cdot 11 \cdot 31 \cdot 37 \cdot 67$.

The existence of $(69)_1$ and $(71)_1$ was verified by

using computers; $(69)_2$ and $(69)_3$ were derived by means of certain graphs. For a more detailed account of these simple groups, see J. Tits, *Séminaire Bourbaki* (1970), No. 375, and the references [18, 19, 20.]

$(72)_1$: A. Rudvalis, $g = 2^{14} \cdot 3^3 \cdot 5^3 \cdot 7 \cdot 13 \cdot 19$, a transitive extension of Tits's simple group, i.e., a normal subgroup of $F'_4(2)$ of index 2. Concerning this group $(72)_1$, see the article of J. H. Conway and D. B. Wales, *J. Algebra*, 27 (1973).

$(73)_1$: M. O'Nan, $g = 2^9 \cdot 3^4 \cdot 5 \cdot 7^3 \cdot 11 \cdot 19 \cdot 31$. This group $(73)_1$ was discovered by O'Nan and the existence was verified by C. Sims, using a computer.

$(73)_2$: B. Fischer, $g = 2^{41} \cdot 3^{13} \cdot 5^6 \cdot 7^2 \cdot 11 \cdot 13 \cdot 19 \cdot 23 \cdot 31 \cdot 47$.

$(73)_3$: B. Fischer and R. Griess, $g = 2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71$.

$(74)_1$: J. G. Thompson, $g = 2^{15} \cdot 3^{10} \cdot 5^3 \cdot 7^2 \cdot 13 \cdot 19 \cdot 31$.

$(74)_2$: K. Harada, $g = 2^{14} \cdot 3^6 \cdot 5^6 \cdot 7 \cdot 11 \cdot 19$.

The existence of $(73)_2$ was suggested by B. Fischer, and then that of $(73)_3$ by B. Fischer and R. Griess. Shortly after this, the existence of $(74)_1$ and $(74)_2$ was suggested by J. G. Thompson, and then Thompson and Harada proved the existence of these groups with the aid of P. Smith and S. Norton, using computers. The existence of $(73)_2$ was established in 1976 by Leon and Sims, using a computer. Very recently (July 1980), R. Griess has announced that the group $(73)_3$ is realized as a group of automorphisms of a 196,883-dimensional commutative nonassociative algebra over the rational numbers.

$(75)_1$: Z. Janko, $g = 2^{21} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11^3 \cdot 23 \cdot 29 \cdot 31 \cdot 37 \cdot 43$. This group $(75)_1$ was discovered by Z. Janko. The existence was verified by using a computer.

For a more detailed account of the twenty-six sporadic simple groups, see [21].

Simple groups of order < 1000 are A_5 ($g = 60$), $L_2(7)$ (168), $A_6(360)$, $L_2(8)$ (504), and $L_2(11)$ (660). All simple groups of order $\leq 20,000$ are known.

Among the known simple groups, the following multiply transitive permutation representations are known: Alternating groups A_n (degree n), A_8 and A_7 (degree 15), A_6 (degree 10), A_5 (degree 6), Mathieu groups M_i (degree i), M_{11} (degree 12), $L_n(q)$ (degree $(q^n - 1)/(q - 1)$), $L_2(p)$ (degree p for $p = 5, 7, 11$), $U_3(q)$ (degree $1 + q^3$), $Sz(q)$ (degree $1 + q^2$), $Re(3^n)$ (degree $1 + 3^{3n}$), $S_{2n}(2)$ (degrees $2^{n-1}(2^n \pm 1)$), and the Higman-Sims group $(67)_2$ (degree 176) and, the Conway group $(68)_4$ (degree 276). Among them, A_n (degree n , $n \geq 5$), M_i (degree i), M_{11} (degree 12), $L_2(2^m)$ (degree $1 + 2^m$), and $L_2(5)$ (degree 5) are triply transitive.

There are several remarkable properties of known finite simple groups which have been conjectured to hold for arbitrary finite simple groups. One of the most famous is the **Schreier conjecture**, which asserts that the group of outer automorphisms of a simple group is solvable. This has been verified for all known cases. Another conjecture says that a finite simple group is generated by two elements. This has also been verified for almost all known groups. In many cases, there is a generating set of two elements, one of which has order 2. There is no counterexample known to disprove the universal validity of this property. Except for $Sz(q)$, the orders of known simple groups are divisible by 12.

J. Classification of Finite Simple Groups

The objective of classification theory is to find the complete list of finite simple groups; this was accomplished in February 1982, following the series of important works mentioned below.

The order of a finite non-Abelian simple group is divisible by at least three distinct prime numbers (W. S. Burnside; \rightarrow Section D). The order of a finite non-Abelian simple group is even (W. Feit and J. G. Thompson; \rightarrow Section D). These theorems are special cases of the following theorem: If G is a finite non-Abelian simple group in which the normalizer of any solvable subgroup $\neq \{1\}$ is solvable, then $G = L_2(q)$ ($q > 3$), $Sz(2^{2n+1})$ ($n \geq 1$), A_7 , $L_3(3)$, $U_3(3)$, M_{11} , or Tits's simple group. In particular, a minimal simple group is isomorphic to $L_2(p)$ ($p \equiv 2$ or $3 \pmod{5}$, $p > 3$), $L_2(2^p)$, $L_2(3^p)$, $Sz(2^p)$, or $L_3(3)$, where p is a prime number (a finite non-Abelian simple group is called a minimal simple group if all proper subgroups are solvable). This theorem is proved in a series of papers by J. G. Thompson (*Bull. Amer. Math. Soc.*, 74 (1968), *Pacific J. Math.*, 33 (1970), 39 (1971), 48 (1973), 50 (1974), and 51 (1974)). The method which Thompson used in these papers has since been generalized in various ways by many authors to establish a number of important theorems. There are also some interesting consequences of this theorem concerning solvable groups. For example, a finite group is solvable if and only if every pair of elements generates a solvable subgroup.

Let G be a non-Abelian simple group of even order and S one of its 2-Sylow subgroups. Then S is neither cyclic nor a generalized quaternion group (W. S. Burnside, R. Brauer, and M. Suzuki). If S is a dihedral group, then $G = A_7$ or $L_2(q)$ (q odd ≥ 5) (D. Gorenstein and J. H. Walter). These theorems deal with the

cases where S is "small." The study in this direction has culminated in the classification of finite simple groups all of whose 2-subgroups are generated by at most four elements (D. Gorenstein and K. Harada, *Mem. Amer. Math. Soc.*, 147 (1974)).

If a 2-Sylow subgroup of a finite non-Abelian simple group G is an Abelian group, then $G = L_2(r)$ ($r \equiv 0, 3$ or $5 \pmod{8}$, $r > 3$), or else G possesses an element of order 2 whose centralizer is isomorphic to $Z_2 \times L_2(q)$ ($q \equiv 3$ or $5 \pmod{8}$, $q > 3$) (J. H. Walter). In the latter case, G is called a **group of Janko-Ree type** (J-R type for short), and if $q \neq 5$, it is called a **group of Ree type**. If $q = 5$, then $G = (64)_1$, the Janko's simple group of order 175,560. The Ree groups $Re(q)$ are groups of Ree type. Since the discovery of $Re(q)$, it has long been an open problem to show that there are no other groups of Ree type. Very recently, the combined work of E. Bombieri, J. G. Thompson and others settled the problem (*Inventiones Math.*, 58 (1980)).

A surprisingly short proof of the Walter's theorem above is given by H. Bender (*Math. Z.*, 117 (1970)). Bender's method applies to a much larger class of groups. A subgroup A of a finite group G is said to be **strongly closed** if $A^g \cap N_G(A) \subset A$ for each $g \in G$. D. M. Goldschmidt proved (*Ann. Math.*, 99 (1974)) that if A is a strongly closed Abelian 2-subgroup, then the subgroup G_0 generated by the conjugates of A possesses a normal series $G_0 \supset G_1 \supset G_2$ with the properties: G_2 is of odd order, G_1/G_2 is a 2-group and is contained in the center of G_0/G_2 , and either $G_0 = G_1$ or G_0/G_1 is the direct product of simple groups on the following list: $L_2(q)$ ($q \equiv 0, 3$ or $5 \pmod{8}$, $q > 3$), $Sz(2^{2n-1})$, $U_3(2^n)$ ($n > 1$), and the groups of J-R type. Furthermore, $AG_2/G_2 \supset G_1/G_2$ and AG_1/G_1 is the center of a 2-Sylow subgroup of G_0/G_1 . This theorem generalizes an earlier result of G. Glauberman (the so-called Z^* -theorem) which states that if A is a strongly closed subgroup of order 2, then the image of A in the quotient group G/K by the maximal normal subgroup of odd order is a normal subgroup and hence is contained in the center of G/K . These theorems of Glauberman and Goldschmidt are of fundamental importance in the study of finite simple groups since they provide an effective tool for showing that a given group has a normal Abelian 2-subgroup.

Glauberman obtained a criterion for the existence of a strongly closed Abelian p -subgroup $\neq \{1\}$ for some prime p [15, 22]. For any prime p , the quadratic group $Qd(p)$ is defined to be the semidirect product of the 2-dimensional vector space $V(2, p)$ over the field of p elements by the special linear group

$SL(2, p)$ where the action of $SL(2, p)$ on $V(2, p)$ is taken to be the natural one. A finite group G contains a strongly closed Abelian p -subgroup $A \neq \{1\}$ if no section of G is isomorphic to $Qd(p)$ (a section is a quotient group of a subgroup). Furthermore, if p is odd, then we can choose as A a characteristic subgroup of a p -Sylow subgroup S of G which is determined only by the structure of S . Therefore a finite non-Abelian simple group has a section isomorphic to $Qd(2) = S_4$ except when it is one of the simple groups mentioned in Goldschmidt's theorem. This theorem generalizes an unpublished result of J. G. Thompson to the effect that 3 divides the order of finite non-Abelian simple groups except $Sz(2^{2n+1})$.

Let G be a 2-transitive permutation group on $n+1$ letters, and assume that the stabilizer H of a letter contains a normal subgroup K which is regular on the remaining n letters. Then G contains a normal subgroup N such that G is isomorphic to a subgroup of the automorphism group of N and either $N = L_2(q)$, $U_3(q)$, $Sz(q)$ or a group of Ree type, or else N is 2-transitive on the $n+1$ letters and no nonidentity element of N leaves two distinct letters invariant (the structure of N in the latter case is also known (H. Zassenhaus; — Section H). This theorem is proved by E. Shult for n even (*Illinois J. Math.*, 16 (1972)) and by C. Hering, W. M. Kantor, and G. M. Seitz for n odd (*J. Algebra*, 20 (1972)). Its proof depends on the work of many authors who considered various special cases, especially the work of H. Zassenhaus, W. Feit, N. Ito, and M. Suzuki on the classification of Zassenhaus groups, and the work of M. Suzuki on the case where n is even and H/K is of odd order (*Ann. Math.*, 79 (1964)). In this special case which Suzuki handles, the stabilizer H is of even order and $H \cap H^g$ is of odd order for any $g \in G - H$. If a proper subgroup H of an arbitrary finite group G has this property, then H is called a **strongly embedded subgroup**. Extending the work of Suzuki, H. Bender proved the following theorem (*J. Algebra*, 17 (1971)): If a finite group G possesses a strongly embedded subgroup, then either (i) a 2-Sylow subgroup of G is cyclic or a generalized quaternion group, or (ii) G possesses a normal series $G = G_0 \supset G_1 \supset G_2$ such that G_0/G_1 and G_2 are of odd order and $G_1/G_2 = L_2(2^n)$, $U_3(2^n)$, or $Sz(2^{2n-1})$ ($n > 1$). This theorem generalizes another theorem of Suzuki who reached the same conclusion under the assumption that two distinct 2-Sylow subgroups have only the identity element in common. Bender's theorem is of fundamental importance in the classification theory of finite simple groups, since a strongly embedded subgroup often appears as an obstacle to the proofs of classification

theorems. For a generalization of Bender's theorem, see a paper by M. Aschbacher (*Proc. Amer. Math. Soc.*, 38 (1973)), which also contains an alternative proof of Shult's theorem.

The theorem of Shult, Hering, Kantor, and Seitz may be interpreted as a classification of finite groups having a split (B, N) -pair of rank 1. Let G be a finite group and let B and N be subgroups of G such that (i) B and N generate G , (ii) $T = B \cap N$ is a normal subgroup of N , and (iii) $W = N/T$ is generated by a set S of elements of order 2 such that $sBs \neq B$ and $sBw \subset BwB \cup BswB$ for each $s \in S$ and each $w \in W$. The subgroups B and N are called a (B, N) -pair of G (the quadruplet (G, B, N, S) is called a **Tits system**), and the cardinality of the set S is called the rank of the (B, N) -pair. The (B, N) -pair is said to be split if B has a normal subgroup U such that $B = TU$ and $T \cap U = \{1\}$, and is said to be saturated if $T = \bigcap_{n \in N} B^n$. If a finite group G has a split saturated (B, N) -pair of rank 1, and if $Z = \bigcap_{g \in G} B^g$, then G/Z is a 2-transitive permutation group satisfying the assumption of the theorem of Shult, Hering, Kantor, and Seitz, and information is obtained on the structure of G . In general, the simple groups of Lie type are characterized as simple groups with certain (B, N) -pairs. For (B, N) -pairs of rank 2, see papers by P. Fong and G. M. Seitz (*Inventiones Math.*, 21 (1973), 24 (1974)). J. Tits has developed a satisfactory theory on finite groups having a (B, N) -pair of rank at least 3 (Lecture notes in math. 386, Springer).

Let G be a finite group generated by a conjugate class D of elements of order 2, and let π be the set of positive integers consisting of the orders of the products of two distinct elements of D . Furthermore, assume that G has no nontrivial solvable normal subgroup. B. Fischer proved (*Inventiones Math.*, 13 (1971)) that if $\pi = \{2, 3\}$, then G contains a normal subgroup isomorphic to one of the following groups: A_n , $S_{2n}(2)$, $O_{2n}(\pm 1, 2)$, $O_n(\pm 1, 3)$, $U_n(2)$, and the three Fischer's simple groups $(68)_5$, $(69)_2$, $(69)_3$. For a generalization of this theorem, see papers by M. Aschbacher (*Math. Z.*, 127 (1972), *J. Algebra*, 26 (1973)). The most powerful result in this direction is given by F. Timmesfeld (*J. Algebra*, 33 (1975), 35 (1975)): Suppose π consists of 2, 4, and odd positive integers. Furthermore, assume that if d and e are in D and de is of order 4, then $(de)^2 \in D$. Then $G = A_6$, $U_3(3)$, the Hall-Janko group $(67)_1$, $L_n(q)$ ($n \geq 3$), $O_{2n+1}(q)$ ($n \geq 3$), $O_{2n}(\pm 1, q)$ ($n \geq 4$), $G_2(q)$, ${}^3D_4(q)$, $F_4(q)$, ${}^2E_6(q)$, $E_6(q)$, $E_7(q)$, or $E_8(q)$, where $q = 2^m$.

Let G be a simple group of even order and let H be the centralizer of an element of order 2. Then the order of G is bounded by a function f of the order h of H (for example, we can

choose $f(h) = \{h(h+1)\}!$; R. Brauer and K. A. Fowler, *Ann. Math.*, 62 (1955)). In particular, there exist only finitely many isomorphism classes of finite simple groups which contain an element of order 2 with a given centralizer H . This fact is a ground for Brauer's program of studying simple groups of even order in terms of the structure of the centralizers of elements of order 2. There are a number of important results concerning Brauer's program [20]. For example, nine sporadic simple groups were discovered in related works. Since 1973, Brauer's program has been improved greatly by M. Aschbacher, D. Gorenstein, and others, and the classification of finite simple groups was finally completed in February 1982 [21, 23].

References

- [1] L. E. Dickson, *Linear groups with an exposition of the Galois field theory*, Teubner, 1901 (Dover, 1958).
- [2] W. Burnside, *Theory of groups of finite order*, Cambridge Univ. Press, second edition, 1911.
- [3] H. Zassenhaus, *Lehrbuch der Gruppentheorie*, Teubner, 1937; English translation, *The theory of groups*, Chelsea, 1958.
- [4] A. Speiser, *Die Theorie der Gruppen von endlicher Ordnung*, Springer, third edition, 1937.
- [5] W. Specht, *Gruppentheorie*, Springer, 1956.
- [6] M. Suzuki, *Structure of a group and the structure of its lattice of subgroups*, *Erg. Math.*, Springer, 1956.
- [7] H. S. M. Coxeter and W. O. J. Moser, *Generators and relations for discrete groups*, *Erg. Math.*, Springer, 1957.
- [8] M. Hall, *The theory of groups*, Macmillan, 1959.
- [9] C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras*, Interscience, 1962.
- [10] W. R. Scott, *Group theory*, Prentice-Hall, 1964.
- [11] M. Hall and J. K. Senior, *The groups of order 2^n ($n \leq 6$)*, Macmillan, 1964.
- [12] H. Wielandt, *Finite permutation groups*, Academic Press, 1964.
- [13] J. Dieudonné, *Sur les groupes classiques*, *Actualités Sci. Ind.*, Hermann, 1948.
- [14] J. Dieudonné, *La géométrie des groupes classiques*, *Erg. Math.*, Springer, 1955.
- [15] D. Gorenstein, *Finite groups*, Harper & Row, 1968. Second edition, Chelsea, 1980.
- [16] D. Passman, *Permutation groups*, Benjamin, 1968.

- [17] B. Huppert, *Endliche Gruppen*, Springer, 1967.
- [18] R. Brauer and C. H. Sah (eds.), *Theory of finite groups*, a symposium, Benjamin, 1969.
- [19] M. B. Powell and G. Higman, (eds.), *Finite simple groups*, Academic Press, 1971.
- [20] W. Feit, The current situation in the theory of finite simple groups, *Actes Congr. Intern. Math.*, 1970, Nice, Gauthier-Villars, vol. 1., p. 55–93.
- [21] D. Gorenstein, The classification of finite simple groups I, *Bull. Amer. Math. Soc.*, 1 (1979), 43–199.
- [22] G. Glaubermann, Factorization in local subgroups of finite groups, *Regional conference series in mathematics*, no. 33 (1977).
- [23] D. Gorenstein, *Finite simple groups*, Plenum, 1982.

152 (VII.18) Finsler Spaces

A. Definitions

Let $T(M)$ be the \dagger tangent vector bundle of an n -dimensional \dagger differentiable manifold M . An element of $T(M)$ is denoted by (x, y) , where x is a point of M and y is a \dagger tangent vector of M at x . Given a \dagger local coordinate system (x^1, \dots, x^n) of M , we can obtain a local coordinate system of $T(M)$ by regarding $(x^1, \dots, x^n, y^1, \dots, y^n) = (x^i, y^j)$ as coordinates of the pair $(x, y) \in T(M)$, where (x^1, \dots, x^n) are coordinates of a point x of M and $y = \sum y^j \partial / \partial x^j$. A continuous real-valued function $L(x, y)$ defined on $T(M)$ is called a **Finsler metric** if the following conditions are satisfied: (i) $L(x, y)$ is differentiable at $y \neq 0$; (ii) $L(x, \lambda y) = |\lambda| L(x, y)$ for any element (x, y) of $T(M)$ and any real number λ ; and (iii) if we put $g_{ij}(x, y) = (1/2) \partial^2 L(x, y)^2 / \partial y^i \partial y^j$, the symmetric matrix $(g_{ij}(x, y))$ is positive definite. A differentiable manifold with a Finsler metric is called a **Finsler space**. There exists a Finsler metric on a manifold M if and only if M is \dagger paracompact. We call $F(x, y) = L(x, y)^2$ the **fundamental form** of the Finsler space. When $F(x, y)$ is a quadratic form of (y^1, \dots, y^n) , $L(x, y)$ is a \dagger Riemannian metric, and $F(x, y) = \sum_{i,j} g_{ij}(x) y^i y^j$. Therefore a Finsler space is a Riemannian space if and only if g_{ij} does not depend on y . The matrix g_{ij} is also called the **fundamental tensor** of the Finsler space $(i, j = 1, \dots, n)$.

Thus the notion of a Finsler metric is an extension of that of Riemannian metric. The study of differentiable manifolds utilizing such generalized metrics was considered by B. Riemann, but he stated that a Riemannian

metric is more convenient for the purpose since “only nongeometrical results can be obtained” by using Finsler metrics [7]. P. Finsler initiated the systematic study of Finsler metrics and extended to a Finsler space many concepts and theorems valid in the classical theory of curves and surfaces [5].

B. The Finsler Metric

In a Finsler space, the arc length of a curve $x = x(t)$ ($a \leq t \leq b$) is given by $\int_a^b L(x, dx/dt) dt$. Therefore a \dagger geodesic in a Finsler space is defined as a \dagger stationary curve for the problem of \dagger variation $\delta \int_a^b L(x, dx/dt) dt = 0$, and the differential equation of the geodesic is given by

$$\frac{d^2 x^i}{dt^2} + \sum_{j,k} \gamma_{jk}^i \left(x, \frac{dx}{dt} \right) \frac{dx^j}{dt} \frac{dx^k}{dt} = 0,$$

where $\gamma_{jk}^i(x, y)$ is the \dagger Christoffel symbol of g_{ij} , i.e.,

$$\gamma_{jk}^i(x, y) = \frac{1}{2} \sum_a g^{ia} \left(\frac{\partial g_{ja}}{\partial x^k} + \frac{\partial g_{ak}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^a} \right),$$

where $(g^{ij}(x, y))$ is the inverse matrix of $(g_{ij}(x, y))$.

The distance between two points in a Finsler space is defined, as in a Riemannian space, as the infimum of the lengths of curves joining the two points. Many properties of Riemannian spaces as metric spaces can be extended to Finsler spaces. The topology defined by the Finsler metric coincides with the original topology of the manifold. A Finsler space M is said to be \dagger complete if every Cauchy sequence of M as a metric space is convergent. The following three conditions are equivalent: (i) M is complete; (ii) each bounded closed subset of M is compact; (iii) each geodesic in M is infinitely extendable. In a complete Finsler space, any two points can be joined by the shortest geodesic.

A diffeomorphism φ of a Finsler space M preserves the distance between an arbitrary pair of points if and only if the transformation on $T(M)$ induced by φ preserves the Finsler metric $L(x, y)$. Such a transformation is called an \dagger isometry of the Finsler space. In the \dagger compact-open topology the set of all isometries of a Finsler space is a \dagger Lie transformation group of dimension at most $n(n+1)/2$. If a Finsler space admits the isometry group of dimension greater than $(n(n-1)+2)/2$, it is a Riemannian space of constant curvature [8].

C. The Theory of Connections

An important difference between a Finsler space and a Riemannian space relates to their

properties with respect to the theory of \dagger connections. In the case of a Riemannian space, the Christoffel symbols constructed from the fundamental tensor are exactly the coefficients of a connection, whereas in the case of a Finsler space, the Christoffel symbols $\gamma_{jk}^i(x, y)$ do not define a connection, for the fundamental tensor g_{ij} depends not only on the points of the space but also on the directions of tangent vectors at these points.

When we consider notions such as tensors, etc., in a Finsler space M , it is generally more convenient to take the whole tangent vector bundle $T(M)$ into consideration rather than restricting ourselves to the space M . For example, let P be the \dagger tangent n -frame bundle over a Finsler space M and $Q = p^{-1}(P)$ be the \dagger principal fiber bundle over $T(M)$ induced from P by the projection p of $T(M)$ onto M . We call the elements of fiber bundles associated with Q tensors. In this sense, the fundamental tensor g_{ij} in a Finsler space is the covariant tensor field of order 2. Therefore it is natural to consider a connection in a Finsler space as a connection in the principal fiber bundle Q . The connection in a Finsler space defined by E. Cartan is exactly of this type [3]. Namely, he showed that by assigning to a connection in Q certain conditions related to the Finsler metric, we can determine uniquely a connection from the fundamental tensor so that the covariant differential of the fundamental tensor vanishes.

Cartan's introduction of the notion of connection produced a development in the theory of Finsler spaces that parallels the development in the theory of Riemannian spaces, and many important results have been obtained. O. Varga (1941) succeeded in obtaining a Cartan connection in a simpler way by using the notion of osculating Riemannian space. S. S. Chern (1943) studied general Euclidean connections that contain Cartan connections as a special case. Noticing that the tangent space of a Finsler space is a \dagger normed linear space, H. Rund (1950) obtained many notions different from those of Cartan. However, as far as the theory of connections is concerned, the two theories do not seem to be essentially different. The theory of curvature in a Finsler space is more complicated than that in a Riemannian space because we have three curvature tensors in the Cartan connection. Using the fact that, in a local cross section of the tangent vector bundle of a Finsler space, a Riemannian metric can be introduced by the Finsler metric, L. Auslander (1955) [2] extended to Finsler spaces the results of J. L. Synge and S. B. Myers on the curvature and topology of Riemannian spaces. A. Lichnérowicz extended the \dagger Gauss-Bonnet formula to

Finsler spaces by considering an integral on the subbundle of the tangent vector bundle satisfying $L(x, y) = 1$ [6]. M. H. Akbar-Zadeh studied \dagger holonomy groups and transformation groups of Finsler spaces by using the theory of fiber bundles.

Connections of Finsler spaces have been investigated by many geometers, but most of them used methods considerably different from those of the modern theory of connections in principal fiber bundles. J. H. Taylor and Synge (1925) defined the covariant differential of a vector field along a curve. L. Berwald (1926) defined a connection from the point of view of the **general geometry of paths**. A curve on a manifold satisfying the differential equation

$$\frac{d^2 x^i}{dt^2} + 2G^i\left(x, \frac{dx}{dt}\right) = 0$$

is called a **path**. The theory was originated by O. Veblen and T. Y. Thomas and generalized as above by J. Douglas. Characteristically, with respect to a Berwald connection, the covariant differential of the fundamental tensor does not vanish.

A Finsler space is a space endowed with a metric for line elements. As a dual concept, we have a **Cartan space**, which is endowed with a metric for **areal elements** [4]. A. Kawaguchi (1937) extended these notions further and studied a **space of line elements of higher order** (or **Kawaguchi space**).

References

- [1] M. H. Akbar-Zadeh, Les espaces de Finsler et certaines de leurs généralisations, *Ann. Sci. Ecole Norm. Sup.*, (3) 80 (1963), 1–79.
- [2] L. Auslander, On curvature in Finsler geometry, *Trans. Amer. Math. Soc.*, 79 (1955), 378–388.
- [3] E. Cartan, *Les espaces de Finsler*, Actualités Sci. Ind., Hermann, 1934.
- [4] E. Cartan, *Les espaces métriques fondés sur la notion d'aire*, Actualités Sci. Ind., Hermann, 1933.
- [5] P. Finsler, *Über Kurven und Flächen in allgemeinen Räumen*, dissertation, Göttingen, 1918.
- [6] A. Lichnérowicz, Quelques théorèmes de géométrie différentielle globale, *Comment. Math. Helv.*, 22 (1949), 271–301.
- [7] B. Riemann, *Über die Hypothesen, welche der Geometrie zu Grunde liegen*, Habilitationsschrift, 1854. (Gesammelte mathematische Werke, Teubner, 1876, 254–270; Dover, 1953.)
- [8] H. C. Wang, On Finsler spaces with completely integrable equations of Killing, *J. London Math. Soc.*, 22 (1947), 5–9.

[9] H. Rund, *The differential geometry of Finsler spaces*, Springer, 1959.

[10] M. Matsumoto, *The theory of Finsler connections*, Publ. Study Group Geom. 5., Okayama Univ., 1970.

153 (IX.7) Fixed-Point Theorems

A. General Remarks

Given a mapping f of a space X into itself, a point x of X is called a **fixed point** of f if $f(x) = x$. When X is a topological space and f is a continuous mapping, we have various theorems concerning the fixed points of f .

B. Fixed-Point Theorems for Polyhedra

(1) Brouwer Fixed-Point Theorem. Let X be a \dagger simplex and $f: X \rightarrow X$ a continuous mapping. Then f has a fixed point in X (*Math. Ann.*, 69 (1910), 71 (1912)).

(2) Lefschetz Fixed-Point Theorem. Let $H_p(X)$ be the p -dimensional \dagger homology group of a \dagger finite polyhedron X (with integral coefficients), $T_p(X)$ the \dagger torsion subgroup of $H_p(X)$, and $B_p(X) = H_p(X)/T_p(X)$. The continuous mapping $f: X \rightarrow X$ naturally induces a homomorphism f_* of the free \mathbb{Z} -module $B_p(X)$ into itself. Let α_p be the \dagger trace of f_* and $\Lambda_f = \sum_{p=0}^n (-1)^p \alpha_p$ ($n = \dim X$). We call this integer Λ_f the **Lefschetz number** of f .

We have the **Lefschetz fixed-point theorem**:

(i) Let f, g be continuous mapping sending X into itself. If f, g are \dagger homotopic ($f \simeq g$), then $\Lambda_f = \Lambda_g$. (ii) If $\Lambda_f \neq 0$, then f has at least one fixed point in X (*Trans. Amer. Math. Soc.*, 28 (1926)).

The condition $\Lambda_f \neq 0$ is, however, not necessary for the existence of a fixed point of f . The Brouwer fixed-point theorem is obtained immediately from (i) and (ii). In particular, if the mapping f is homotopic to the identity mapping 1_X , then α_p is the p th \dagger Betti number of X , and Λ_f is equal to the \dagger Euler characteristic $\chi(X)$ of X . Hence, in this case, if $\chi(X) \neq 0$, then f has a fixed point.

When X is a compact oriented manifold without boundary the Lefschetz number Λ_f of f can be interpreted as the \dagger intersection number of the graph of f and the diagonal of X .

More generally, let X and Y be compact oriented n -dimensional manifolds without boundary. If f and $g: X \rightarrow Y$ are continuous mappings, a point x of X such that $f(x) =$

$g(x)$ is called a **coincidence point** of f and g . The intersection number $\Lambda_{f,g}$ of the graph of f and that of g is called the **coincidence number** of f and g . If $\Lambda_{f,g} \neq 0$, then f and g have at least one coincidence point. The coincidence number $\Lambda_{f,g}$ is also expressed as $\sum_{p=0}^n (-1)^p \text{tr}(f^* \circ g_! | H^p(X))$, where $g_!: H^p(X) \rightarrow H^p(Y)$ is the \dagger Gysin homomorphism of g .

Suppose that a finite group G acts on the manifolds X and Y . If $f: X \rightarrow Y$ is a mapping, a point x of X such that $\tau f(x) = f\tau(x)$ for all $\tau \in G$ is called an **equivariant point** of f . When G is a group of order 2 and acts on X nontrivially, the **equivariant point index** $\hat{\Lambda}_f$ is employed. This index was introduced by Nakaoka (*Japan. J. Math.* 4 (1978)), using the \dagger equivariant cohomology. It has the property that $\hat{\Lambda}_f \neq 0$, implying that f admits an equivariant point. The prototype of this theorem is the Borsuk-Ulam theorem (*Fund. Math.* 20 (1933)), which states that a continuous mapping $f: S^n \rightarrow \mathbb{R}^n$ always admits a point $x \in S^n$ such that $f(x) = f(-x)$.

(3) Lefschetz Number and Fixed-Point Indices. Suppose that $|K|$ is an n -dimensional homogeneous polyhedron (i.e., any simplex of K that is not a face of another simplex of K is of dimension n), and $f: |K| \rightarrow |K|$ is a continuous mapping. Then there exists a continuous mapping $g: |K| \rightarrow |K|$ homotopic to f and admitting only isolated fixed points $\{q_1, \dots, q_r\}$, each of which is an inner point of an n -dimensional simplex of K . The \dagger local degree λ_i of a mapping g at q_i is called the **fixed-point index** of g at q_i . Then $J_f = \sum_{i=1}^r \lambda_i$ does not depend on the choice of g and is equal to $(-1)^n \Lambda_f$.

(4) Singularities of a Continuous Vector Field.

Let X be an n -dimensional \dagger differentiable manifold and F a \dagger continuous vector field on X that assigns a tangent vector x_p to each point p of X . A point p is called a **singular point** of F if x_p is the zero vector. The vector field F induces in a natural manner a continuous mapping $f: X \rightarrow X$ that is homotopic to the identity mapping 1_X . Then a fixed point of f is a singular point of F , and vice versa. When such a singular point p is isolated, there exists a \dagger coordinate neighborhood N of p that is homeomorphic to an n -dimensional open ball such that x_q is nonzero for every point q in N except for $q = p$. Let N' be the boundary of N . Then we may consider $N' \cong S^{n-1}$ and a mapping $F|N'$ from S^{n-1} to $\mathbb{R}^n \setminus \{0\} \simeq S^{n-1}$. The \dagger degree of a mapping $S^{n-1} \xrightarrow{F|N'} \mathbb{R}^n \setminus \{0\} \xrightarrow{\cong} S^{n-1}$ is called the **index of the singular point** p . This index is equal to the fixed-point index λ_p of f at p . Hence, when X is compact and has no

boundary, the sum of indices of (isolated) singular points of F is equal to $(-1)^n \chi(X)$. In particular, a compact manifold X without boundary admits a continuous vector field with no singular point if and only if $\chi(X) = 0$ (Hopf's theorem, *Math. Ann.*, 96 (1927)).

(5) Poincaré-Birkhoff Fixed-Point Theorem. In certain cases, a continuous mapping $f: X \rightarrow X$ of a finite polyhedron X into itself has fixed points even if $\Lambda_f = 0$. For example, let X be the annular space $\{(r, \theta) | \alpha \leq r \leq \beta\}$ ((r, θ) are the polar coordinates of points in a Euclidean plane) and let $f: X \rightarrow X$ be a homeomorphism satisfying the following conditions: (i) there exist continuous functions $g(\theta)$, $h(\theta)$ such that $g(\theta) < \theta$, $h(\theta) > \theta$, $f(\alpha, \theta) = (\alpha, g(\theta))$, $f(\beta, \theta) = (\beta, h(\theta))$; (ii) there exists a continuous positive function $\rho(r, \theta)$ defined for $\alpha < r < \beta$ such that

$$\iint_D \rho(r, \theta) dr d\theta = \iint_D \rho(f(r, \theta)) dr d\theta$$

for all measurable sets D .

Then f has at least two fixed points. This theorem was conjectured in 1912 by Poincaré, who hoped to apply it to solve the \dagger restricted three-body problem. The theorem was later proved by G. D. Birkhoff (*Trans. Amer. Math. Soc.*, 14 (1913); see also M. Brown and W. D. Neumann, *Michigan Math. J.* 24 (1977)) and is called the **Poincaré-Birkhoff fixed-point theorem** or the **last theorem of Poincaré**.

C. Atiyah-Bott and Atiyah-Singer Fixed-Point Theorems

There is a far-reaching generalization of the Lefschetz formula given by Atiyah and Bott (*Ann. Math.*, (2) 86 (1967), 88 (1968)). Let M be a compact differentiable manifold without boundary and $f: M \rightarrow M$ a differentiable mapping with only simple fixed points; that is, it is assumed that $\det(1 - df_p) \neq 0$ for each fixed point $p \in M$ of f , where df_p is the differential of f at the point p . The fixed points of f are finite in number. Suppose that an \dagger elliptic complex over M

$$\mathcal{E}: 0 \rightarrow \Gamma(E_0) \xrightarrow{d_0} \Gamma(E_1) \xrightarrow{d_1} \dots \xrightarrow{d_{l-1}} \Gamma(E_l) \rightarrow 0$$

and a sequence of smooth vector bundle homomorphisms $\varphi_i: f^*E_i \rightarrow E_i$ ($i = 0, \dots, l$) are given such that $d_i T_i = T_{i+1} d_i$ for each i , where $T_i: \Gamma(E_i) \rightarrow \Gamma(E_i)$ is defined by $T_i s(x) = \varphi_i s(f(x))$ for $s \in \Gamma(E_i)$. The sequence $T = (T_i)$ induces endomorphisms $H^i T$ of the homology groups $H^i(\mathcal{E})$ of the elliptic complex \mathcal{E} . We define the Lefschetz number $L(T)$ by $L(T) = \sum_{i=0}^l (-1)^i \text{tr } H^i T$. On the other hand, for a fixed point p of f , let $\varphi_{i,p}: E_{i,p} \rightarrow E_{i,p}$ denote the restriction of φ_i on the fiber $E_{i,p}$ of E_i over

p . Under the circumstances mentioned above the Lefschetz number $L(T)$ is given by the formula

$$L(T) = \sum_p \sum_{i=0}^l \frac{(-1)^i \text{tr } \varphi_{i,p}}{|\det(1 - df_p)|},$$

where the summation is over the fixed points of f .

Here are some examples of the above formula. First, take as \mathcal{E} the \dagger de Rham complex and as φ_i the obvious one, i.e., the i th exterior power of the transpose of df . In this case, the formula reduces to the classical Lefschetz formula. As a second example consider the \dagger Dolbeault complex

$$0 \rightarrow \Lambda^{0,0}(M) \xrightarrow{\bar{\partial}} \Lambda^{0,1}(M) \xrightarrow{\bar{\partial}} \dots \xrightarrow{\bar{\partial}} \Lambda^{0,n}(M) \rightarrow 0$$

of a compact complex manifold M and a holomorphic mapping $f: M \rightarrow M$ with only simple fixed points. In this case the formula above reduces to one giving the Lefschetz number of the induced endomorphisms $f^*: H^{0,*}(M) \rightarrow H^{0,*}(M)$ of the \dagger Dolbeault cohomology:

$$L(f^*) = \sum_p \frac{1}{\det_C(1 - df_p)},$$

where df_p is regarded as a holomorphic differential.

If the assumption that the mapping f has only simple fixed points is replaced by the one that f is a diffeomorphism of M contained in a compact \dagger transformation group G , then there is also a generalized Lefschetz formula, given by Atiyah and Singer (*Ann. Math.*, (2) 87 (1968)). The fixed-point set of such a diffeomorphism is a closed submanifold of M (consisting of several components). Suppose that we are given an elliptic complex \mathcal{E} over M and a lift of the G -action on M to \mathcal{E} . The latter implies that, if we define $T_i: \Gamma(E_i) \rightarrow \Gamma(E_i)$ by $T_i s(x) = f^{-1} s(f(x))$ for $s \in \Gamma(E_i)$, then $d_i T_i = T_{i+1} d_i$ holds for each i . Under these circumstances, the Lefschetz number $L(T)$ can be expressed in the form

$$L(T) = \sum_{F_i} v(F_i),$$

where the summation is over the components $\{F_i\}$ of the fixed-point set M^f of f and where the number $v(F_i)$ is written explicitly in terms of the \dagger symbol of the elliptic complex \mathcal{E} with G -action, the characteristic classes of the manifold F_i , the characteristic classes of the normal bundle of F_i in M , and the action of $g = f^{-1}$ on the normal vectors. The formula is essentially a reformulation of the \dagger Atiyah-Singer index theorem. In fact, $L(T)$ is the \dagger analytic index of \mathcal{E} evaluated on g and the number $v(F_i)$ is deduced from the \dagger topological index of \mathcal{E} using

the localization theorem. The most useful elliptic complexes are de Rham complexes, Dolbeault complexes, signature operators and Dirac operators. In the case of Dolbeault complexes, f is assumed to be an analytic automorphism, and the number $v(F_i)$ takes the form

$$v(F_i) = \frac{\prod_0 \mathcal{Q}^\theta \mathcal{T}(F_i)}{\det(1 - g|N^*)} [F_i].$$

Here, the normal bundle N of F_i has a decomposition $N = \sum_\theta N(\theta)$ into the sum of complex vector bundles such that g acts on $N(\theta)$ as $e^{i\theta}$, and \mathcal{Q}^θ is the characteristic class defined by

$$\mathcal{Q}^\theta = \prod_j \frac{1 - e^{-i\theta}}{1 - e^{-x_j - i\theta}},$$

where the \dagger Chern class of $N(\theta)$ is written as $c(N(\theta)) = \prod_j (1 + x_j)$; moreover $\mathcal{T}(F_i)$ denotes the \dagger Todd class of the complex manifold F_i , and N^* denotes the dual bundle of N (\rightarrow 237 K-Theory H).

D. Fixed-Point Theorems for Infinite-Dimensional Spaces

Birkhoff and O. D. Kellogg generalized Brouwer's fixed-point theorem to the case of function spaces (*Trans. Amer. Math. Soc.*, 23 (1922)). Their result was utilized to show the existence of solutions of certain differential equations, and has led to a new method in the theory of functional equations.

J. P. Schauder obtained the following theorem: Let A be a closed convex subset of a Banach space, and assume that there exists a continuous mapping T sending A to a \dagger countably compact subset $T(A)$ of A . Then T has fixed points (*Studia Math.*, 2 (1930)). This theorem is called the **Schauder fixed-point theorem**.

A. Tikhonov generalized Brouwer's result and obtained the following **Tikhonov fixed-point theorem** (*Math. Ann.*, 111 (1935)): Let R be a locally convex \dagger topological linear space, A a compact convex subset of R , and T a continuous mapping sending A into itself. Then T has fixed points.

This theorem may be applied to the case where R is the space of continuous mappings sending an m -dimensional Euclidean space E^m into a k -dimensional Euclidean space E^k to show the existence of solutions of certain differential equations. For example, when $m = k = 1$, consider the differential equation

$$dy/dx = f(x, y), \quad y(x_0) = y_0.$$

We set $T(y) = y_0 + \int_{x_0}^x f(t, y(t)) dt$ to determine a continuous mapping $T: R \rightarrow R$. Then the fixed points of T are the solutions of the differential

154 A Foliations

equation. Now we can apply the theorem of Tikhonov to show the existence of solutions.

On the other hand, when we are given problems of functional analysis, Schauder's fixed-point theorem is usually more convenient to apply than Tikhonov's theorem.

The following theorem, written in terms of functional analysis, is useful for applications: Let D be a subset of an n -dimensional Euclidean space, F the family of continuous functions defined on D , and $T: F \rightarrow F$ a mapping. Suppose that the following three conditions are satisfied: (i) For $f_1, f_2 \in F$, $0 < \lambda < 1$ implies $\lambda f_1 + (1 - \lambda)f_2 \in F$. (ii) If a series $\{f_k\}$ of functions in F converges uniformly in the wider sense to a function f , then $f \in F$; and furthermore, the series $\{Tf_k\}$ converges uniformly in the wider sense to Tf . (iii) The family $T(F)$ is a \dagger normal family of functions on D . Then there exists a function $f \in F$ such that $Tf = f$.

Let R be a topological linear space and T a mapping assigning a closed convex subset $T(x)$ of R to each point x of R . A point x of R is called a **fixed point** of T if $x \in T(x)$. The mapping T is called **semicontinuous** if the condition $x_n \rightarrow a$, $y_n \rightarrow b$ ($y_n \in T(x_n)$) implies that $b \in T(a)$. In particular, if K is a bounded closed convex subset of a finite-dimensional Euclidean space R and T a semicontinuous mapping sending points of K into convex subsets of K , then T admits fixed points (**Kakutani fixed-point theorem**, *Duke Math. J.*, 8 (1941)). This result was further generalized to the case of locally convex topological linear spaces by Ky Fan (*Proc. Nat. Acad. Sci. US*, 38 (1952)).

References

- [1] B. F. Brown, The Lefschetz fixed point theorem, Scott, Foresman, 1971.
- [2] J. Cronin, Fixed points and topological degree in nonlinear Analysis, Amer. Math. Soc. Math. Surveys, no. 11, 1964.
- [3] S. Lefschetz, Topology, Amer. Math. Soc. Colloq. Publ., vol. XII, 1930.
- [4] D. R. Smart, Fixed point theorems, Cambridge tracts in math. 66, Cambridge Univ. Press, 1974.

154 (IX.21) Foliations

A. Introduction

A foliation is a kind of geometric structure on manifolds, such as a differentiable or complex structure. The study of foliations evolved from

investigation of the behavior of orbits of a vector field and also of the solutions of total differential equations. Through the early works of C. Ehresmann, G. Reeb, and A. Haefliger, together with the development of manifold theory in 1960s, it became an established field of mathematics. Since then, great progress has been made in this field, especially in its topological aspects. At the same time it turns out that many problems in foliation theory are deeply related not only to the geometry of manifolds but also to various other branches of mathematics, such as the theory of differential equations, functional analysis, and group theory.

B. Definitions and General Remarks

A **foliation** on a manifold can be defined within various categories: topological, C^r -differentiable ($1 \leq r \leq \infty$), real analytic, and holomorphic. For definiteness, however, we restrict ourselves to the C^r -differentiable category in what follows. Furthermore, all manifolds are assumed to be paracompact.

Let M be an n -dimensional C^∞ -manifold, possibly with boundary. A **codimension** q , C^r -**foliation** of M ($0 \leq q \leq n, 0 \leq r \leq \infty$) is a family $\mathcal{F} = \{L_\alpha \mid \alpha \in A\}$ of arcwise connected subsets, called **leaves**, of M with the following properties: (i) $L_\alpha \cap L_{\alpha'} = \emptyset$ if $\alpha \neq \alpha'$; (ii) $\bigcup_{\alpha \in A} L_\alpha = M$; (iii) Every point in M has a local coordinate system (U, ψ) of class C^r such that for each leaf L_α the arcwise connected components of $U \cap L_\alpha$ are described by $x^{n-q+1} = \text{constant}, \dots, x^n = \text{constant}$, where x^1, x^2, \dots, x^n denote the local coordinates in the system (U, ψ) . In particular, every leaf of \mathcal{F} is an $(n-q)$ -dimensional submanifold of M . The totality of integral submanifolds of a completely integrable nonsingular system of Pfaffian equations on \mathbf{R}^n , $\omega_i = a_{i1}(x)dx_1 + a_{i2}(x)dx_2 + \dots + a_{in}(x)dx_n = 0$ ($i = 1, 2, \dots, q$) forms a codimension q foliation, and the totality of integral curves of a nonsingular vector field of class C^r on M ($r \geq 1$) constitutes a codimension $n-1$ C^r -foliation.

Let \mathcal{F} be a codimension q , C^r -foliation of M ($r \geq 1$). Then M admits a C^{r-1} p -plane field consisting of all vectors tangent to the leaves of \mathcal{F} , and, dually, a C^{r-1} q -plane field ($p+q=n$). Denote the former by $\tau(\mathcal{F})$ and the latter by $\nu(\mathcal{F})$, and call them the **tangent bundle** and the **normal bundle** of \mathcal{F} , respectively. $\nu(\mathcal{F})$ is isomorphic to the quotient bundle $T(M)/\tau(\mathcal{F})$. \mathcal{F} is called **transversely orientable** if $\nu(\mathcal{F})$ is an orientable vector bundle. A C^r -mapping $f: N \rightarrow M$ is said to be **transverse to the foliation** \mathcal{F} if the composite mapping $T(N) \xrightarrow{df} T(M) \rightarrow T(M)/\tau(\mathcal{F})$ is epimorphic at each point of N .

In this case, f induces a codimension q , C^r -foliation $f^{-1}(\mathcal{F})$ of N whose leaves are the arcwise connected components of $f^{-1}(L)$ ($L \in \mathcal{F}$).

In particular, if Q is a q -dimensional C^r -manifold and $f: N \rightarrow Q$ is a C^r -submersion, f induces a codimension q , C^r -foliation of N whose leaves are the arcwise connected components of $f^{-1}(x)$ ($x \in Q$).

A C^r p -plane field \mathcal{X} on M is called **involutive** if, for any C^r vector fields X, Y on M such that $X_x, Y_x \in \mathcal{X}_x$ ($x \in M$), the Lie bracket $[X, Y]$ satisfies $[X, Y]_x \in \mathcal{X}_x$. This condition is known as the **Frobenius integrability condition** for \mathcal{X} . If \mathcal{X} is defined locally by q Pfaffian equations $\omega_1 = \dots = \omega_q = 0$, the above condition is equivalent to the condition that there are C^{r-1} 1-forms θ_{ij} ($i, j = 1, \dots, q$) such that $d\omega_i = \sum_{j=1}^q \theta_{ij} \wedge \omega_j$. A C^r p -plane field \mathcal{X} is said to be **completely integrable** if it is a tangent bundle of some foliation. When $r \geq 2$, \mathcal{X} is completely integrable if and only if it is involutive (**Frobenius theorem**) (\rightarrow 428 Total Differential Equations). There is a topological obstruction to the complete integrability of \mathcal{X} (\rightarrow Section F).

A closed C^∞ -manifold M admits a codimension 1, C^r -plane field if and only if the Euler number of M vanishes. In 1944, Reeb constructed a codimension 1, C^∞ -foliation of the 3-sphere S^3 as follows [1]. Let $f(x)$ be a C^∞ -even function defined on the open interval $(-1, 1)$, such that

$$\lim_{|x| \rightarrow 1} \frac{d^{(k)}}{dx^{(k)}} \frac{1}{f'(x)} = 0 \quad (k = 0, 1, 2, \dots).$$

The graphs of the equations $y = f(x) + c$ ($-1 < x < 1, c \in \mathbf{R}$) together with the lines $x = \pm 1$ constitute a codimension 1, C^∞ -foliation of $[-1, 1] \times \mathbf{R}$. Then by rotating it around the y -axis in \mathbf{R}^3 , we obtain a codimension 1, C^∞ -foliation of $D^2 \times \mathbf{R}$, where D^2 denotes the closed 2-disk. The foliation is invariant under vertical translations and therefore defines a codimension 1, C^∞ -foliation of $D^2 \times S^1$. This foliation is called the **Reeb component** of $D^2 \times S^1$. Since S^3 is a union of two solid tori intersecting in the common toroidal surface, the Reeb components of each solid torus constructed above define the so-called **Reeb foliation** of S^3 (Fig. 1).

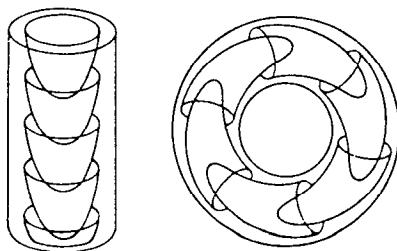


Fig. 1

Generalizing the above construction into a differential topological method, one obtains the following results: Every closed 3-dimensional manifold admits a codimension 1, C^∞ -foliation (S. P. Novikov [3], W. Lickorish, J. Wood); every odd-dimensional sphere admits a codimension 1, C^∞ -foliation (I. Tamura, A. Durfee, B. Lawson). On the other hand, every open manifold has a codimension 1, C^∞ -foliation induced by a submersion over \mathbf{R} (\rightarrow Section F).

Let M be a total space of a C^r -bundle over a C^∞ -manifold B with fiber F . If F is a C^∞ -manifold and the \dagger structure group reduces to a totally disconnected subgroup of $\text{Diff}^r F$, the group of all C^r -diffeomorphisms of F , then the local sections, which are defined in an obvious manner using the local triviality of this bundle, fit together to give leaves of a codimension q , C^r -foliation ($q = \dim F$). In this case, M is called a **foliated bundle** or a **flat F -bundle** over B . Each leaf of this foliation is diffeomorphic to a covering manifold of B and transverse to the fibers of the bundle $M \rightarrow B$. Foliated bundles exhibit a class of foliations; this is especially important in connection with the characteristic classes of foliations (\rightarrow Section G).

C. Holonomy

The notion of the holonomy of a leaf, given by Ehresmann, is a generalization of the \dagger Poincaré mapping in \dagger dynamical systems. Let \mathcal{F} be a codimension q , C^r -foliation of M and L be a leaf of \mathcal{F} . Let $N(L)$ denote the total space of the normal disk bundle of L in M . Choose a C^r -immersion $i: N(L) \rightarrow M$ such that i restricted to the zero section of $N(L)$ is the natural inclusion and i maps the fibers of $N(L)$ transversely to the foliation \mathcal{F} . Then the induced foliation $i^{-1}(\mathcal{F})$ of $N(L)$ has the properties that the leaves are transverse to the fibers of $N(L)$ and the zero section of $N(L)$ is a leaf. If γ is an oriented loop in L based at a point $x_0 \in L$, then there is a neighborhood U of 0 in the fiber over x_0 satisfying the following: for each point $x \in U$ there is a curve $\gamma_x: [0, 1] \rightarrow N(L)$ having the properties: (i) $\gamma_x(0) = x$, (ii) $\text{Im}(\gamma_x)$ lies on a leaf of $i^{-1}(\mathcal{F})$, and (iii) $\pi \circ \gamma_x(t) = \gamma(t)$ for any $t \in [0, 1]$, where $\pi: N(L) \rightarrow L$ is the bundle projection. The family of curves $\{\gamma_x | x \in U\}$ gives a C^r -diffeomorphism H_γ from U to another open set of $\pi^{-1}(x_0)$, which assigns $\gamma_x(1)$ to x . Let G_q^r denote the group of \dagger germs at 0 of all local C^r -diffeomorphisms of \mathbf{R}^q which fix 0. The germ at 0 of the mapping H_γ depends only on the \dagger homotopy class of γ , and, by identifying $\pi^{-1}(x_0)$ with \mathbf{R}^q , we obtain a homomorphism $h_L: \pi_1(L, x_0) \rightarrow G_q^r$. h_L is determined by L up to conjugacy and is called the **holonomy homo-**

morphism, or simply the **holonomy**, of the leaf L . The image of h_L is called the **holonomy group** of L . For $r \geq 1$, by differentiating each element of G_q^r , one has a homomorphism $dh_L: \pi_1(L, x_0) \rightarrow GL(q; \mathbf{R})$, called the **linear holonomy** of L . The holonomy of a proper leaf (\rightarrow Section D) completely characterizes the foliation of a neighborhood of it (Haefliger).

D. Topology of Leaves

Let \mathcal{F} be a codimension q foliation of M . The **leaf topology** is a topology of M defined by requiring each connected component of the set of the form $U \cap L$ to be open, where U is an open set in M and $L \in \mathcal{F}$. Leaves are nothing but the connected components of M with respect to this topology. A leaf $L \in \mathcal{F}$ is called a **compact leaf** if L is compact in the leaf topology. In general, L is called **proper** if two topologies on L induced from the original and the leaf topologies on M coincide. Any compact leaf is proper. A leaf L is called **locally dense** if $\text{Int } \bar{L} \neq \emptyset$. If a leaf is neither proper nor locally dense, it is called **exceptional**. Since we are assuming that M is paracompact, a leaf that is a closed subset of M is proper (Haefliger). There exists a codimension 1, C^1 -foliation of the 2-torus T^2 that contains exceptional leaves. But in the C^r category ($r \geq 2$), such a foliation does not exist on T^2 (A. Denjoy, C. Siegel). In higher dimensions, there are examples of C^∞ -foliations with exceptional leaves (R. Sacksteder). The following result is called the **Novikov closed leaf theorem** [3]: Any codimension 1, C^r -foliation ($r \geq 2$) of a closed 3-dimensional manifold M contains a Reeb component if either $\pi_1(M)$ is finite or $\pi_2(M) \neq 0$ ($M \neq S^1 \times S^2, S^1 \times RP^2$). In particular, every C^2 -foliation of S^3 contains a compact leaf homeomorphic to T^2 . The question of whether every codimension 2, C^r -foliation of S^3 has a compact leaf is known as the **Seifert conjecture**. There is a counterexample in the C^1 case (P. Schweitzer [7]), but it remains unsolved for $r \geq 2$ (\rightarrow 126 Dynamical Systems N).

A compact leaf L is said to be **stable** if it has an arbitrarily small open neighborhood that is a union of compact leaves. The following results are called the **Reeb stability theorems**: (1) Let L be a compact leaf of a C^r -foliation ($r \geq 0$). If the holonomy group of L is finite, then it is stable (Reeb [1]). (2) Let \mathcal{F} be a transversely orientable codimension 1, C^r -foliation ($r \geq 1$) of a compact manifold M (tangent to the boundary). If there exists a compact leaf L with $H^1(L; \mathbf{R}) = 0$, then M is a fiber bundle over S^1 or $[0, 1]$, and the leaves of \mathcal{F} are the fibers of this bundle. In particular,

every leaf is compact and stable (Reeb [1], Thurston [8]).

The generalization of the stability theorem for proper leaves has been investigated by T. Inaba and P. Dippolito.

E. Haefliger Structures

Let \mathcal{G}_q be the \dagger pseudogroup consisting of all C^r -diffeomorphisms g of an open subset of \mathbf{R}^q to another open subset of \mathbf{R}^q . We write Γ_q^r for the set of all germs $[g]_x$ of g at x , $x \in \text{domain of } g$, $g \in \mathcal{G}_q$. The sheaf topology of Γ_q^r is the topology whose open base is the family of subsets of the form $\bigcup_{x \in \text{domain of } g} \{[g]_x\}$. With this topology and the multiplication induced from the composition of \mathcal{G}_q , Γ_q^r is a **topological groupoid**, i.e., a \dagger groupoid whose multiplication and inverse mappings are continuous. If one identifies a point x in \mathbf{R}^q with $[\text{id}_{\mathbf{R}^q}]_x$, Γ_q^r contains \mathbf{R}^q as a subspace.

A codimension q , C^r -**Haefliger structure** or a Γ_q^r -**structure** \mathcal{H} on a topological space X is a maximal covering of X by open sets $\{U_i | i \in J\}$, such that for each pair $i, j \in J$, there is a continuous mapping $\gamma_{ij}: U_i \cap U_j \rightarrow \Gamma_q^r$ satisfying

$$\gamma_{ik}(x) = \gamma_{ij}(x) \circ \gamma_{jk}(x) \text{ for } x \in U_i \cap U_j \cap U_k. \quad (*)$$

Since $\gamma_{ii}(x)$ is the germ of the identity mapping for $x \in U_i$, the correspondence $x \rightarrow \gamma_{ii}(x)$ defines a continuous mapping $f_i: U_i \rightarrow \mathbf{R}^q \subset \Gamma_q^r$. A codimension q , C^r -foliation of a C^∞ -manifold M is the same as a Γ_q^r -structure on M such that each f_i is a C^r -submersion (Haefliger [5]).

If $f: Y \rightarrow X$ is continuous and \mathcal{H} is a Γ_q^r -structure on X , there is an induced Γ_q^r -structure $f^{-1}\mathcal{H}$ on Y which is defined by $\{f^{-1}(U_i), \gamma_{ij} \circ f | i, j \in J\}$. Two Γ_q^r -structures \mathcal{H}_0 and \mathcal{H}_1 on X are said to be **homotopic** if there exists a Γ_q^r -structure \mathcal{H} on $X \times [0, 1]$ such that $\mathcal{H}|_{X \times \{t\}} = \mathcal{H}_t$ ($t = 0, 1$).

Let $\Gamma_q^r(X)$ be the set of homotopy classes of Γ_q^r -structures on X . There exists a space $B\Gamma_q^r$, called the **classifying space for Γ_q^r -structures**, such that there is a natural one-to-one correspondence between $\Gamma_q^r(X)$ and $[X, B\Gamma_q^r]$ for any paracompact space X , where $[A, B]$ denotes the set of homotopy classes of continuous mappings from A to B . By condition (*) above, if $r > 0$, the differentials $\{d\gamma_{ij}(x) | x \in U_i \cap U_j, i, j \in J\}$ define a q -dimensional vector bundle $\nu(\mathcal{H})$ over X , which is called the **normal bundle** of \mathcal{H} . The correspondence $\mathcal{H} \rightarrow \nu(\mathcal{H})$ gives a continuous mapping $\nu: B\Gamma_q^r \rightarrow BGL(q, \mathbf{R})$ among classifying spaces. If $r = 0$, there is also a similar mapping $\nu: B\Gamma_q^0 \rightarrow B\text{Top}_q$. Let $B\Gamma_q^0$ be the homotopy fiber of the mapping ν . $B\Gamma_q^0$ is a classifying space for the Γ_q^0 -structures with trivialized normal bundles.

There is a tight connection between the

topology of $B\Gamma_q^r$ and the group structure of $\text{Diff}^r(\mathbf{R}^q)$, which is stated below. Let $\text{Diff}_K^r(\mathbf{R}^q)$ be the topological group of all C^r -diffeomorphisms of \mathbf{R}^q that are identities outside some compact sets with the C^r topology. Let $\text{Diff}_{K,\delta}^r(\mathbf{R}^q)$ denote $\text{Diff}_K^r(\mathbf{R}^q)$ with the discrete topology and $B\overline{\text{Diff}}_K^r(\mathbf{R}^q)$ denote the homotopy fiber of the natural mapping $B\text{Diff}_{K,\delta}^r(\mathbf{R}^q) \rightarrow B\text{Diff}_K^r(\mathbf{R}^q)$. Then there exists a continuous mapping $B\overline{\text{Diff}}_K^r(\mathbf{R}^q) \rightarrow \Omega^q B\Gamma_q^r$ that induces an isomorphism in the homology group with integer coefficient for $0 \leq r \leq \infty$ and $q \geq 1$, where Ω^q denotes the q th \dagger loop space functor (J. N. Mather [12], Thurston [10]). Further, it has been proved that $\text{Diff}_{K,\delta}^r(\mathbf{R}^q)$ is a \dagger simple group (if $r \neq q + 1$) and that $B\Gamma_q^r$ is $\dagger(q + 1)$ -connected for $r \neq q + 1$ (Mather and Thurston, Haefliger [5]). The group $\text{Homeo}_{K,\delta}\mathbf{R}^q$ is \dagger acyclic (Mather), and hence $B\Gamma_q^0$ is \dagger contractible.

F. Existence and Classification of Foliations

Not every plane field on a manifold is isomorphic to a completely integrable field (R. Bott). Thus, in general, the existence of a foliation of a manifold. Haefliger and Thurston solved the existence and classification problems in foliation in terms of Haefliger structures as follows. Two codimension q foliations \mathcal{F}_0 and \mathcal{F}_1 of a C^∞ -manifold M are said to be **concordant** if there is a codimension q foliation \mathcal{F} on $M \times [0, 1]$, that is transverse to $M \times \{t\}$ ($t = 0, 1$) and induces there the given foliation \mathcal{F}_t ($t = 0, 1$). They are said to be **integrably homotopic** if one further requires that the foliation \mathcal{F} be transverse to $M \times \{t\}$ for all $t \in [0, 1]$ in the definition above. Similarly, two subbundles ξ_0, ξ_1 of $T(M)$ are said to be **concordant** if there is a subbundle ξ of $T(M \times [0, 1])$ such that $\xi|_{M \times \{t\}} = \xi_t$ for $t = 0, 1$, and they are said to be **homotopic** if one further assumes that $\xi|_{M \times \{t\}}$ is a subbundle of $T(M \times \{t\})$ for all $t \in [0, 1]$. The following theorem is of fundamental importance.

Theorem: Let M be an open (resp. closed) C^∞ -manifold. Then for each $r = 0, 1, \dots, \infty$, the integrable homotopy classes (resp. concordance classes) of codimension q , C^r -foliations of M are in a natural one-to-one correspondence with homotopy classes of Γ_q^r -structures \mathcal{H} on M together with homotopy classes (resp. concordance classes) of subbundles of $T(M)$ isomorphic to $\nu(\mathcal{H})$. (M. L. Gromov, A. Phillips, Haefliger [5], Thurston [9]).

The following are consequences of the theorem: (i) A closed manifold M admits a codimension 1, C^∞ -foliation if and only if the Euler number of M vanishes. (ii) If a manifold

admits a ${}^{\dagger}q$ -frame field, then the associated q -plane field is homotopic to the normal bundle of a codimension q foliation of M . (iii) Every dimension q plane field on a C^{∞} -manifold is homotopic to the normal bundle of a C^0 -foliation with C^{∞} leaves.

G. Characteristic Classes of Foliations

Let $B\Gamma_q^r$ be the classifying space of Γ_q^r -structures. An element of the cohomology group $H^*(B\Gamma_q^r; \mathbf{R})$ is called a (real) **characteristic class of codimension q , C^r -foliations**. If \mathcal{F} is a codimension q , C^r -foliation of M and $f: M \rightarrow B\Gamma_q^r$ is the classifying mapping for \mathcal{F} , then an element $\alpha(\mathcal{F}) = f^* \alpha \in H^*(M; \mathbf{R})$, $\alpha \in H^*(B\Gamma_q^r; \mathbf{R})$, is called the characteristic class of \mathcal{F} corresponding to α . The first nontrivial characteristic classes of foliations are known as the **Godbillon-Vey classes** (C. Godbillon and J. Vey [13]) and can be defined as follows. Let \mathcal{F} be a transversely oriented, codimension q , C^{∞} -foliation of M . Then there exists a ${}^{\dagger}q$ -form Ω on M such that on a neighborhood U of each point of M , Ω is written as $\omega_1^U \wedge \dots \wedge \omega_q^U$, where $\omega_1^U, \dots, \omega_q^U$ are linearly independent 1-forms that vanish on leaves of \mathcal{F} . By the integrability condition, there is a 1-form η such that $d\Omega = \eta \wedge \Omega$. Then the Godbillon-Vey class $\Gamma_{\mathcal{F}}$ of \mathcal{F} is the de Rham cohomology class in $H^{2q+1}(M; \mathbf{R})$ represented by the closed $2q+1$ form $\eta \wedge (d\eta)^q$.

The following construction provides a wide class of characteristic classes of foliations (Bott and Haefliger [14], I. Bernstein and B. Rosenfeld [15]). Let J_k be the set of ${}^{\dagger}k$ -jets at 0 of local C^{∞} -diffeomorphisms of \mathbf{R}^q keeping 0 fixed. The set $\{J_k\}_{k=0}^{\infty}$ forms an † inverse system of † Lie groups with respect to the natural homomorphism $p_k: J_{k+1} \rightarrow J_k$, and each J_k ($k \geq 1$) contains $O(q)$ as a † maximal compact subgroup. Let P_k be the differentiable fiber bundle of k -jets at 0 of local diffeomorphisms of \mathbf{R}^q whose domains contain 0. It is a † principal J_k -bundle over \mathbf{R}^q . Denote by $A(P_{\infty})$ the † direct limit of the de Rham complexes of $\{P_k\}_{k=0}^{\infty}$, and let A be the subcomplex of $A(P_{\infty})$ consisting of invariant forms with respect to the natural action of \mathcal{G}_q^{∞} . A is canonically isomorphic to the † cochain complex $A(A_q)$ of continuous alternating forms on A_q , where A_q is the topological Lie algebra of † formal vector fields on \mathbf{R}^q (\rightarrow 105 Differentiable Manifolds AA).

Now let \mathcal{F} be a codimension q , C^{∞} -foliation of M . Let $P_k(\mathcal{F})$ denote the differentiable fiber bundle over M whose fiber over $x \in M$ is the space of k -jets at x of the C^{∞} -submersion $f: U \rightarrow \mathbf{R}^q$ from an open neighborhood U of x to \mathbf{R}^q , satisfying (i) $f(x) = 0$, (ii)

$\text{Ker } df = \tau(\mathcal{F})|_U$. This is a principal J_k -bundle over M , and its restriction to U is isomorphic to the pullback by f of the bundle P_k . Hence there are homomorphisms from the set of invariant forms on P_k to $A(P_k(\mathcal{F}))$ that induce a homomorphism $A \cong A(A_q) \rightarrow \varinjlim A(P_k(\mathcal{F}))$. This homomorphism is compatible with the action of $O(q)$ and hence induces a homomorphism of $O(q)$ -basic subcomplexes. Thus one obtains a homomorphism $\varphi_{\mathcal{F}}: H^*(A_q; O(q)) \rightarrow \varinjlim H^*(A(P_k(\mathcal{F})); O(q)) \cong H^*(M; \mathbf{R})$. In fact, $\varphi_{\mathcal{F}}$ depends only on the 2-jet of the foliation \mathcal{F} , and one can think of it as a homomorphism $\varphi: H^*(A_q; O(q)) \rightarrow H^*(B\Gamma_q^r; \mathbf{R})$ ($r \geq 2$). The elements in $\text{Im } \varphi$ are called the **smooth characteristic classes of foliations**.

Let WO_q be the differential graded algebra:

$$WO_q = E(u_1, u_3, \dots, u_{2[(q+1)/2]-1}) \otimes \hat{\mathbf{R}}[c_1, \dots, c_q],$$

where $du_i = c_i$, $dc_i = 0$, $\deg(u_i) = 2i - 1$, $\deg(c_i) = 2i$, and E denotes the exterior algebra over \mathbf{R} generated by the u_i 's, and $\hat{\mathbf{R}}$ denotes the † polynomial algebra of the c_i 's truncated by the † ideal generated by elements of degree $> 2q$. There exists a homomorphism of differential graded algebras $WO_q \rightarrow A(A_q; O(q))$ which induces an isomorphism in cohomology (I. M. Gel'fand and D. B. Fuks). For a codimension q foliation \mathcal{F} of M , the cohomology class determined by c_{2i} corresponds to the i th Pontryagin class of the normal bundle $\nu(\mathcal{F})$ of \mathcal{F} , and the cohomology class $u_1 c_q^q$ corresponds to the Godbillon-Vey class of \mathcal{F} . In particular, the subring of the cohomology ring $H^*(M; \mathbf{R})$ generated by the † Pontryagin classes of $\nu(\mathcal{F})$ is trivial for degree $> 2q$ (Bott's vanishing theorem).

Let N^{q+1} be a closed $(q+1)$ -dimensional Riemannian manifold of † constant negative curvature. The total space $T_1 N$ of the unit tangent sphere bundle of N admits a codimension q , C^{∞} -foliation associated with the † geodesic flow of $T_1 N$ († Anosov foliation). It has been shown that the Godbillon-Vey class of this foliation is nontrivial (R. Roussarie, F. Kamber and P. Tondeur, K. Yamato). It is known that many of the smooth characteristic classes are also nontrivial (Bott and Haefliger, Thurston, J. Heitsch).

A smooth characteristic class $\alpha \in H^*(B\Gamma_q^r; \mathbf{R})$ is called **rigid** if for any smooth one-parameter family $\{\mathcal{F}_t\}$ of codimension q foliations on a C^{∞} -manifold M , $d(\alpha(\mathcal{F}_t))/dt = 0$ holds. The elements in the image of the natural homomorphism

$$H^*(WO_{q+1}) \rightarrow H^*(WO_q) \rightarrow H^*(B\Gamma_q^r; \mathbf{R})$$

are rigid (Heitsch). On the other hand, the Godbillon-Vey class is not rigid. In fact, Thurston constructed a one-parameter family of codimension q foliations of a certain

$(2q+1)$ -dimensional manifold for which the Godbillon-Vey class varies continuously. The characteristic classes of a simple foliation are often trivial. For example, the Godbillon-Vey class of a codimension 1 foliation of a closed manifold vanishes if it is almost without holonomy (i.e., no noncompact leaves have non-trivial holonomy) (M. Herman [16]; T. Mizutani, S. Morita, and T. Tsuboi [17]).

H. Further Topics

(1) **Transverse structures.** Let B be a q -dimensional manifold and \mathcal{S} a geometric structure on it, and let $\mathcal{G}_{\mathcal{S}}$ denote the \dagger pseudo-group generated by the local diffeomorphisms that preserve the structure \mathcal{S} . Replacing \mathcal{G}_q by $\mathcal{G}_{\mathcal{S}}$ in the definition of a C^r -Haefliger structure, one obtains definitions of a $\Gamma_{\mathcal{S}}$ -structure and a $\Gamma_{\mathcal{S}}$ -foliation. A $\Gamma_{\mathcal{S}}$ -foliation is called a **Riemannian foliation**, a transversely **real analytic foliation**, or a transversely **holomorphic foliation** if \mathcal{S} is a Riemannian, real analytic, or complex structure on \mathbf{R}^q ($\mathbf{C}^{q/2}$), respectively. The theories for many such foliations are analogous to those for C^r -foliations. For example, many results are known about the characteristic classes of Riemannian or holomorphic foliations. Haefliger showed that there is no codimension 1 real analytic foliation on a simply connected closed manifold and that the classifying space $B\bar{\Gamma}_1^w$ for codimension 1 transversely oriented transversely real analytic foliations has the homotopy type of a $\dagger K(\pi, 1)$ -space for some uncountable \dagger perfect group π [5].

(2) **Foliated cobordism.** Two closed oriented n -dimensional C^∞ -manifolds M_0 and M_1 with codimension q , C^r -foliations are said to be **foliated cobordant** if there exist a compact oriented $(n+1)$ -dimensional C^∞ -manifold W with boundary $\partial W = M_1 \cup (-M_0)$ and a codimension q , C^r -foliation of W which is transverse to ∂W and induces the given foliations of M_0 and M_1 . The resulting foliated cobordism classes $\{\mathcal{F}\}$ form a group $\mathcal{F}\Omega_{n,q}^r$ with respect to the disjoint union. It is known that $\mathcal{F}\Omega_{2,1}^\infty = \{0\}$ and that the Reeb foliation of S^3 is cobordant to zero. The characteristic classes of foliations provide invariants of foliated cobordisms. In particular the Godbillon-Vey number $\Gamma_{\mathcal{F}}[M]$ is an invariant of $\mathcal{F}\Omega_{2q+1,q}^r$ ($r \geq 2, q \geq 1$), and a result of Thurston mentioned in Section G implies that the homomorphism $\mathcal{F}\Omega_{2q+1,q}^r \rightarrow \mathbf{R}$ defined by $\{\mathcal{F}\} \mapsto \Gamma_{\mathcal{F}}[M]$ is surjective.

(3) **Growth of leaves, transverse invariant measure.** Let \mathcal{F} be a C^∞ -foliation of a compact manifold. Fix a \dagger Riemannian metric on M . Then each leaf L of \mathcal{F} has the induced metric, and one has a function $f_L(r) =$

$\text{vol}(D(x, r))$, where $D(x, r)$ is the set of points $y \in L$ whose distance along L from a fixed point $x \in L$ is not greater than r . The growth type of the function f_L is determined only by L . Many papers have been published that deal with the relation between the behavior of leaves and their growth types in codimension 1 foliations (J. Cantwell and L. Conlon, G. Hector, T. Nishimori, N. Tsuchiya). On the other hand, as a generalization of the notion of asymptotic cycles in dynamical systems, the notion of **foliation cycles**, or equivalently a **transverse invariant measure**, has been defined (J. Plante [18], D. Ruelle, D. Sullivan [19]). The existence of a transverse invariant measure for a foliation is closely connected to the growth types of the leaves. The example of Denjoy in Section D leads to the study of \dagger minimal sets of foliations and the structures of foliations (Hector, Cantwell and Conlon). The structure of a codimension 1 foliations which are almost without holonomy has been fairly well investigated (Sacksteder, Hector, H. Imanishi, R. Moussou, Roussaire).

(4) **Compact foliation.** A foliation whose leaves are all compact is called a **compact foliation**. D. Epstein proved that if \mathcal{F} is a codimension 2, C^2 -compact foliation of a closed 3-manifold, then the leaves of \mathcal{F} are the fibers of a \dagger Seifert fibration of M . In higher dimensions, the situation is more complicated (Sullivan, R. Edwards and K. Millett).

(5) **Foliated bundles.** There are many results on foliated bundles. In particular J. W. Milnor [20] and J. Wood obtained a condition for a circle bundle ξ over a closed surface Σ to have a foliation transverse to fibers. More precisely, if ξ and Σ are orientable, then ξ admits such a foliation if and only if $|X(\xi)| \leq -\min\{0, \chi(\Sigma)\}$, where X denotes the Euler number and χ the Euler-Poincaré characteristic. Kamber and Tondeur made an extensive study of characteristic classes of foliated bundles [21].

(6) **Transverse foliations.** Two foliations \mathcal{F} and \mathcal{G} of M are said to be **transverse** to each other if any two leaves K and L of \mathcal{F} and \mathcal{G} are transverse to each other. A foliated bundle has such foliations. D. Hardorp proved that on every orientable closed 3-manifold, there exists a triple of codimension 1 foliations that are pairwise transverse. Tamura and A. Sato classified the codimension 1 foliations that are transverse to the Reeb component of $D^2 \times S^1$.

References

- [1] G. Reeb, Sur certaines propriétés topologiques de variétés feuilletées, *Actualité Sci. Ind.*, 1183, Hermann, 1952.
- [2] C. Ehresmann, Sur la théorie des variétés

- feuilletées, Univ. Roma Rend. Mat. Appl., 10 (1951), 64–82.
- [3] S. P. Novikov, Topology of foliations, Amer. Math. Soc. Transl., 1967, 286–304. (Original in Russian, 1965.)
- [4] I. Tamura, Foliations and spinnable structures on manifolds, Ann. Inst. Fourier, 23 (1973), 197–214.
- [5] A. Haefliger, Homotopy and integrability, Lecture notes in math. 197, Springer, 1971, 133–163.
- [6] A. Denjoy, Sur les courbes définies par les équations différentielles à la surface du tore, J. Math. Pure Appl., 11 (1932), 333–375.
- [7] P. Schweitzer, Counterexamples to Seifert conjecture and opening closed leaves of foliations, Ann. Math., (2) 100 (1974), 386–400.
- [8] W. Thurston, A generalization of Reeb stability theorem, Topology, 13 (1974), 347–352.
- [9] W. Thurston, Existence of codimension one foliations, Ann. Math., (2) 104 (1976), 249–268.
- [10] W. Thurston, Foliations and groups of diffeomorphisms, Bull. Amer. Math. Soc., 80 (1974), 304–307.
- [11] W. Thurston, Noncobordant foliations of S^3 , Bull. Amer. Math. Soc., 78 (1972), 511–514.
- [12] J. N. Mather, Integrability in codimension 1, Comment. Math. Helv., 48 (1973), 195–233.
- [13] C. Godbillon and J. Vey, Un invariant des feuilletages de codimension un, C. R. Acad. Sci. Paris, 273 (1971), 92–95.
- [14] R. Bott and A. Haefliger, On characteristic classes of Γ -foliations, Bull. Amer. Math. Soc., 78 (1972), 1039–1044.
- [15] I. Bernstein and B. Rosenfeld, On characteristic classes of foliations, Functional Anal. Appl., 6 (1972), 60–62.
- [16] M. Herman, The Godbillon-Vey invariant of foliations by planes of T^3 , Lecture notes in math. 597, Springer, 1977, 294–307.
- [17] T. Mizutani, S. Morita, and T. Tsuboi, The Godbillon-Vey classes of codimension one foliations which are almost without holonomy, Ann. Math., (2) 113 (1981).
- [18] J. Plante, Foliations with measure-preserving holonomy, Ann. Math., (2) 102 (1975), 327–362.
- [19] D. Sullivan, Cycles for the dynamical study of foliated manifolds and complex manifolds, Inventiones Math., 36 (1976), 225–255.
- [20] J. W. Milnor, On the existence of a connection with curvature zero, Comment. Math. Helv., 32 (1958), 215–223.
- [21] F. Kamber and P. Tondeur, Foliated bundles and characteristic classes, Lecture notes in math. 493, Springer, 1975.
- [22] B. Lawson, Foliations, Bull. Amer. Math. Soc., 80 (1974), 369–418.

- [23] B. Lawson, The quantitative theory of foliations, Regional conference series in math. 27, Conference Board of the Mathematical Sciences, 1977.

155 (VI.2) Foundations of Geometry

A. Introduction

Geometry deals with figures. It depends, therefore, on our spatial intuition, but our intuition lacks objectivity. The Greeks originated the idea of developing geometry logically, based on explicitly formulated axioms, without resorting to intuition. From this intention resulted Euclid's *Elements*, which was long considered the perfect model of a logical system. As time passed, however, mathematicians came to notice its imperfections. Since the 19th century especially, with the awakening of a more rigorous critical spirit in science and philosophy, more systematic criticism of the *Elements* began to appear. Non-Euclidean geometry was formulated after reexamination of Euclid's axiom of parallels; but it was also discovered that even as a foundation of Euclidean geometry, Euclid's system of axioms was far from perfect. Various systems of axioms for Euclidean geometry were proposed by mathematicians in the latter half of the 19th century, among them one by D. Hilbert [1], which became the basis of far-reaching studies.

B. Hilbert's System of Axioms

Hilbert took as undefined elements **points** (denoted by A, B, C, \dots), **straight lines** (or simply **lines**, denoted by a, b, c, \dots), and **planes** (denoted by $\alpha, \beta, \gamma, \dots$). Between these objects there exist incidence relations (expressed in phrases such as " A lies on a ," " a passes through A ," etc.); order relations (" B is between A and C "); congruence relations; and parallel relations. The relations are subject to the following five groups of axioms:

- (I) **Incidence axioms:** (1) For two points A, B , there exists a line a through A and B . (2) If $A \neq B$, the line a through A, B is uniquely determined. We write $a = A \cup B$ and call a the **join** of A, B . (3) Every line contains at least two different points. There exist at least three points that do not lie on a line. (4) If A, B, C are points not on a line, there exists a plane α through A, B, C . (We also say that A, B, C lie on α .) For every plane α , there exists at least one point A on α . (5) If A, B, C are points

not on a line, the plane α through A, B, C is uniquely determined. We write $\alpha = A \cup B \cup C$ and call α the **join** of A, B, C . (6) If A, B are two different points on a line a and if A, B lie on a plane α , then every point on a lies on α . (We say that a lies on α or α passes through a .) (7) If a point A lies on two planes α, β , there exists at least one other point B on α and β . (8) There exist at least four points not lying on a plane.

(II) **Ordering axioms:** (1) If B is between A and C , then A, B, C are three different points lying on a line; also, B is between C and A . (2) If A, C are two different points, then there exists a point B such that C is between A and B . (3) If B is between A and C , then A is not between B and C .

We define a **segment** as a set of two different points A, B , denoted as AB or BA , and we call A and B **ends** of this segment. The set of points between A, B is called the **interior** of AB , and the set of points of $A \cup B$ that are neither ends nor interior points of AB is called the **exterior** of AB .

(4) Let A, B, C be three points not lying on a line. If a line a on the plane $A \cup B \cup C$ does not pass through A, B , or C , but passes through a point of the interior of AB , then it also passes through a point of the interior of BC or CA (**Pasch's axiom**).

The following propositions are proved from the above axioms. Given n points A_1, A_2, \dots, A_n on a line ($n > 2$), we can rearrange them, if necessary, so that the point A_j is between A_i and A_k whenever we have $1 \leq i < j < k \leq n$. There are exactly two ways of arranging the points in this manner (**theorem of linear ordering**). Let O be a point on a line a , and let A, B be two points on the line different from O . Write $A \sim B$ when $A = B$ or O is not between A and B ; write $A \not\sim B$ otherwise. Then \sim is an equivalence relation between points on the line different from O ; from $A \sim B, A \sim C$ it follows that $B \sim C$. We say that A, B are on the same **side** or on different sides of O on a depending on whether $A \sim B$ or $A \not\sim B$. Two subsets a' and a'' of a defined by $a' = \{A' | A \sim A'\}, a'' = \{A'' | A \not\sim A''\}$ are called **half-lines** or **rays** on a with O as the extremity (or starting from O). Denoting by a , for simplicity, the set of points on a , we have $a = a' \cup \{O\} \cup a''$ (disjoint union).

Using axiom II.4, we can also prove the following: Let a be a line on α , and let A, B be two points on α not lying on a . If $A = B$ or if the interior of the segment AB has no point in common with a , we say that A, B are on the same **side** of a on α , and write $A \sim B$. Otherwise, we say that A, B are on different sides of a on α and write $A \not\sim B$. Then \sim is an equivalence relation between points on α not lying on a , and from $A \sim B, A \sim C$ follows $B \sim C$. The

subsets $\alpha' = \{A' | A \sim A'\}, \alpha'' = \{A'' | A \not\sim A''\}$ of α are called **half-planes** on α **bounded** by a . Again, denoting by α and a the set of points on α and the set on a , respectively, we obtain $\alpha = \alpha' \cup a \cup \alpha''$ (disjoint union).

(III) **Congruence axioms:** Two segments $AB, A'B'$ can be in a relation of **congruence**, expressed symbolically as $AB \equiv A'B'$. (Segments AB and $A'B'$ are then said to be **congruent**. Since the segment AB is defined as the set $\{A, B\}$, the four relations $AB \equiv A'B', BA \equiv A'B', AB \equiv B'A', BA \equiv B'A'$ are equivalent.) This relation is subject to the following three axioms: (1) Let A, B be two different points on a line a , and A_1 a point on a line a_1 (a_1 may or may not be equal to a). Let a'_1 be a ray on a_1 starting from A_1 . Then there exists a unique point B_1 on a'_1 such that $AB \equiv A_1B_1$. (2) From $A_1B_1 \equiv AB$ and $A_2B_2 \equiv AB$ follows $A_1B_1 \equiv A_2B_2$. (Hence it follows that \equiv is an equivalence relation between segments.) (3) Let A, B, C be three points such that B is between A and C , and let A_1, B_1, C_1 be three points such that B_1 is between A_1 and C_1 . Then from $AB \equiv A_1B_1, BC \equiv B_1C_1$ follows $AC \equiv A_1C_1$.

Now let h, k be two different lines in a plane α and through a point O , and let h', k' be the rays on h, k starting from O . The set of two such rays h', k' is called an **angle** in α , denoted by $\angle(h', k')$ or $\angle(k', h')$. This angle is also denoted by $\angle AOB$, where A, B are points of h', k' , respectively. The rays h', k' are called the **sides** and the point O is called the **vertex** of this angle. Then h' is a subset of a half-plane on α bounded by k , and k' is a subset of a half-plane on α bounded by h . The intersection of these two half-planes is called the **interior** of this angle, and the subset of $\alpha - O$ consisting of points belonging to neither the inside nor the sides of the angle is called the **exterior** of the angle. Between two angles $\angle(h', k'), \angle(h'_1, k'_1)$ there may exist the relation of congruence, again expressed by the symbol \equiv , as in the case of segments, and subject to the following two axioms: (4) Let $\angle(h', k')$ be an angle on a plane α and h_1 be a line on α_1 (α_1 may or may not be equal to α). Let O_1 be a point on h_1, h'_1 a ray on α_1 starting from O_1 , and α'_1 a half-plane on α_1 bounded by h_1 . Then there exists a unique ray k'_1 starting from O_1 and lying in α'_1 such that $\angle(h', k') \equiv \angle(h'_1, k'_1)$. Moreover, $\angle(h', k') \equiv \angle(h', k')$ always holds. (Hence it follows that \equiv is an equivalence relation between angles.) (5) Let both A, B, C and A_1, B_1, C_1 be triples of points not lying on a line. Then from $AB \equiv A_1B_1, AC \equiv A_1C_1$, and $\angle BAC \equiv \angle B_1A_1C_1$, it follows that $\angle ABC \equiv \angle A_1B_1C_1$.

(IV) **Axiom of parallels:** Suppose that a, b are two different lines. Then it follows from axiom I.2 that if a and b share a point P , such a point is the unique point lying on both a and

b . In this case we say that a, b **intersect** at P and write $a \cap b = P$. On the other hand, if a and b have no point in common and if a, b are on the same plane, we say that a, b are **parallel** and write $a \parallel b$. If A, a are on a plane α and A is not on a , we can prove (utilizing axioms I, II, and III) that there exists a line b passing through A in α such that $a \parallel b$. The axiom of parallels postulates the uniqueness of such a b .

(V) **Axioms of continuity**: (1) Let AB, CD be two segments. Then there exist a finite number of points A_1, A_2, \dots, A_n on $A \cup B$ such that $CD \equiv AA_1 \equiv A_1A_2 \equiv \dots \equiv A_{n-1}A_n$ and B is between A and A_n (**Archimedes' axiom**). (2) The set of points on a line a (again denoted for simplicity by a) is "maximal" in the following sense: It should satisfy axioms II.1–II.3, III.1, V.1, and the theorem of linear ordering. If \bar{a} is a set of points satisfying these axioms such that $\bar{a} \supset a$, then \bar{a} should be $= a$ (**axiom of linear completeness**). Hence follows the **theorem of completeness**: the set of points, lines, and planes is maximal in the sense that it is not possible to add further points, lines, or planes to this set with the resulting set still satisfying axioms I–IV and V.1.

C. Consistency

In formulating the above axioms and proving their consistency, Hilbert assumed the consistency of the theory of real numbers (\rightarrow 156 Foundations of Mathematics). To prove consistency, Hilbert constructed a model for the above axioms using the method of analytic geometry. He defined points as triples of real numbers (x_1, x_2, x_3) , lines and planes as sets of points satisfying suitable systems of linear equations, and relations of ordering, congruence, and parallelism in the usual way. It is easy to verify that such a system satisfies all the axioms I–V. Thus the consistency of these axioms is reduced to the consistency of the theory of real numbers (\rightarrow 35 Axiom Systems).

A model for I–IV and V.1 can be obtained in the countable field \mathbf{R}_0 of all real \dagger algebraic numbers instead of \mathbf{R} . Then \mathbf{R}_0 can be further restricted to its subfield \mathbf{P}_0 defined as follows: Let F be an arbitrary field. An \dagger extension of F of the form $F(\sqrt{1+\lambda^2})$ with $\lambda \in F$ is called a **Pythagorean extension** of F , and F is said to be a **Pythagorean field** if any Pythagorean extension of F coincides with F (e.g., \mathbf{R}_0 and \mathbf{R} are Pythagorean). It is easily verified that I–IV are satisfied in the "analytic geometry over any Pythagorean field." On the other hand, we can construct a minimal Pythagorean field containing a given field (the **Pythagorean closure** of the field) in the same way as we construct the \dagger algebraic closure of a field. The field \mathbf{P}_0 is

defined as the Pythagorean closure of the field \mathbf{Q} of rational numbers.

D. Independence of Axioms

In Hilbert's system, the axioms I and II are used to formulate further axioms. On the other hand, it can be shown that each of the groups III, IV, and V is independent from other axioms.

The independence of IV is shown by the consistency of non-Euclidean geometry (\rightarrow 285 Non-Euclidean Geometry). The following model shows the independence of III.5: In the analytic model for I–V, we replace the definition of distance between two points

$(x_1, x_2, x_3), (y_1, y_2, y_3)$ by

$$((x_1 - y_1 + x_2 - y_2)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2)^{1/2}.$$

Then III.5 does not hold, while all other axioms remain satisfied. The independence of V.2 is shown by the geometry over \mathbf{R}_0 or \mathbf{P}_0 . The independence of V.1 follows from the existence of the non-Archimedean Pythagorean field: the Pythagorean closure of any \dagger non-Archimedean field (e.g., the field of rational functions of one variable over \mathbf{Q} with a \dagger non-Archimedean valuation) is such a field. A geometry in which V.1 does not hold is called a **non-Archimedean geometry**.

E. Completeness of the System of Axioms and Relations between Axioms

The \dagger completeness of the system of axioms I–V can be shown by introducing coordinates in the geometry with these axioms and representing it as \dagger Euclidean geometry of three dimensions. Axiom group V is essential for the introduction of coordinates over \mathbf{R} . Moreover, we have the following results:

(i) The geometry with the axioms I–IV can be represented as "Euclidean geometry" of three dimensions over a Pythagorean field, and the geometry with axioms I.1–I.3, II–IV can be represented as "Euclidean geometry" of two dimensions over a Pythagorean field.

(ii) The geometry with II, and a stronger axiom of parallels IV* (given a line a and a point A outside a , there exists one and only one line a' passing through A that is parallel to a) can be represented as an \dagger affine geometry over a field K that is not necessarily commutative.

(iii) The field K is commutative if and only if the following holds: Suppose that in Fig. 1 $A' \cup B \parallel A \cup B'$, $B' \cup C \parallel B \cup C'$. Then it follows that $A' \cup C \parallel A \cup C'$ (**Pascal's theorem**).

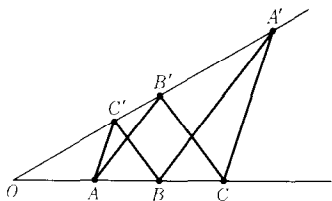


Fig. 1

(iv) The “two-dimensional geometry” with I.1–I.3, II, and IV* can be embedded in the “three-dimensional geometry” with axioms I, II, and IV* if and only if the following holds: Suppose that in Fig. 2 we have $AUB \parallel A'UB'$, $BUC \parallel B'UC'$. Then it follows that $CUA \parallel C'UA'$ (Desargues's theorem).

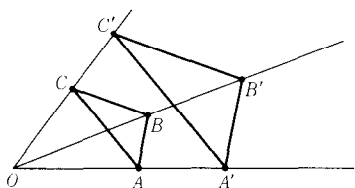


Fig. 2

(v) From I.1–I.3, II, IV*, and Pascal's theorem follows Desargues's theorem.

(vi) Desargues's theorem is independent of I.1–I.3, II, III.1–III.4, IV*, and V; that is, we can construct a **non-Desarguesian geometry** (a geometry in which Desargues's theorem does not hold) in which these axioms are satisfied.

Axioms I, II, and IV*, as well as the theorems of Pascal and Desargues, are propositions in affine geometry. Each has a corresponding proposition in [†]projective geometry, and the results concerning them can be transferred to the case of projective geometry (\rightarrow 343 Projective Geometry).

F. Polygons and Their Areas

Suppose that we are given a finite number of points A_i ($i=0, 1, \dots, r$) in the geometry with axioms I and II. Then the set of segments (or, more precisely, the union of segments together with their interiors) A_iA_{i+1} ($i=0, 1, \dots, r-1$) is called a **broken line** joining A_0 with A_r . In particular, if $A_0 = A_r$, then this set is called a **polygon** with **vertices** A_i and **sides** A_iA_{i+1} . A polygon with r vertices is called an r -gon. (For $r=3, 4, 5, 6$, r -gons are called **triangles**, **quadrangles**, **pentagons**, and **hexagons**, respectively.) A **plane polygon** is a polygon whose vertices all lie on a plane. A polygon is called **simple** if any three consecutive vertices do not lie on a line, and two sides A_iA_{i+1} and A_jA_{j+1} ($i \neq j$) meet only when $j=i+1$ or $i=j+1$. In this article,

we consider only simple plane polygons, and refer to them simply as polygons.

[†]Jordan's theorem implies that a polygon in the sense just defined divides the plane into two parts, its **interior** and its **exterior**. This special case of Jordan's theorem can be proved by I.1–I.3 and II only. A polygon P is divided into two polygons P_1, P_2 by a broken line joining two points on sides of P and lying in the interior of P (Fig. 3). In this case, we say that P is **decomposed** into P_1, P_2 and write $P = P_1 + P_2$. We may again decompose P_1, P_2 and thereby arrive at a decomposition of the form $P = P_1 + \dots + P_k$. Axiom III is used to introduce the congruence relation \equiv between polygons. Two polygons P, Q are called **decomposition-equal** if there exist decompositions $P = P_1 + \dots + P_k, Q = Q_1 + \dots + Q_k$ such that $P_1 \equiv Q_1, \dots, P_k \equiv Q_k$. This is expressed by PzQ . We call P, Q **supplementation-equal** if there exist two polygons P', Q' such that $(P + P')z(Q + Q'), P'zQ'$. This will be expressed by PeQ . If we assume IV, we can use result (i) of Section E. Let K be the ground field of the geometry (K is Pythagorean, hence [†]ordered). The **area** of polygon P is defined as the positive element $m(P)$ of K assigned to P such that $m(P + Q) = m(P) + m(Q)$, and $m(P) = m(P')$ if $P \equiv P'$. From PzQ or PeQ , it follows that $m(P) = m(Q)$. Under these axioms, it is proved that $m(P) = m(Q)$ implies PeQ . If we also assume V.1, then $m(P) = m(Q)$ implies PzQ . Thus the theory of area of polygons can be constructed without assuming axiom V.2, though this result cannot be generalized to higher-dimensional cases. For the case of three dimensions, we can construct two solids of the same volume that are not supplementation-equal [2, 7].

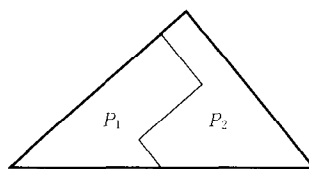


Fig. 3

G. Geometric Construction by Ruler and a Transferrer of Constant Lengths

The geometry with I–IV can be represented as 3-dimensional Euclidean geometry over a Pythagorean field. Conversely, all these axioms are valid in 3-dimensional Euclidean geometry over any Pythagorean field. Thus the minimal system of “quantities” whose existence is assured in geometry with these axioms is the field P_0 , the Pythagorean closure of Q . Hilbert noticed that the existence of a

geometric object under axioms I–IV can be expressed as its constructibility by **ruler** (i.e., an instrument to draw a straight line joining two points) and a **transferrer of constant lengths**. The latter, for a constant length x , is an instrument that permits finding the point X on the given ray AB such that $AX = x$. It is not possible to construct by ruler and transferrer all the points that can be constructed by means of ruler and compass (\rightarrow 179 Geometric Construction). However, it is possible to construct all the lengths λx , where λ is any element of P_0 . Hilbert conjectured that an element of P_0 can be characterized as a †totally positive algebraic number of degree 2^v , $v \in \mathbb{N}$. This conjecture was proved by Artin [3].

H. Related Topics

While Hilbert's foundations are concerned with 3-dimensional Euclidean geometry, it is easy to generalize these results to the case of n -dimensional Euclidean geometry (\rightarrow 139 Euclidean Geometry). Also, for affine and projective geometries, there are well-organized systems of axioms (at least for the case of dimensions ≥ 3). Hilbert [1, Appendix III] showed that plane †hyperbolic geometry can be constructed on a modified system of axioms, but for other non-Euclidean geometries (in particular, telliptic geometries) there are no known systems of axioms as good as Hilbert's for the Euclidean case. On the other hand, Hilbert [1, Appendix IV] gives another method of constructing Euclidean geometry in characterizing the group of motions as the topological group with certain properties. G. Thomsen [4] rewrote Hilbert's system of axioms in group-theoretical language utilizing the fact that the group of motions is generated by symmetries with respect to points, lines, and planes. Finally, Hilbert's study of the foundations of geometry led him to research in the †foundations of mathematics.

References

- [1] D. Hilbert, *Grundlagen der Geometrie*, Teubner, 1899, seventh edition, 1930.
- [2] M. Dehn, Über den Rauminhalt, *Math. Ann.*, 55 (1902), 465–478.
- [3] E. Artin, Über die Zerlegung definiter Funktionen in Quadrate, *Abh. Math. Sem. Univ. Hamburg*, 5 (1926), 100–115 (Collected papers, Addison-Wesley, 1965, 273–288).
- [4] G. Thomsen, *Grundlagen der Elementargeometrie*, Teubner, 1933.
- [5] K. Reidemeister, *Grundlagen der Geometrie*, Springer, 1968.

[6] G. Hessenberg and J. Diller, *Grundlagen der Geometrie*, de Gruyter, 1967.

[7] V. G. Boltyanskii, Hilbert's third problem, Winston, 1978.

156 (I.1) Foundations of Mathematics

A. General Remarks

The notion of †set, introduced toward the end of the 19th century, has proved to be one of the most fundamental and useful ideas in mathematics. Nonetheless, it has given rise to well-known †paradoxes. Based on this notion, R. Dedekind developed the theories of natural numbers [2] and real numbers [3], defining the latter as “cuts” of the set of rational numbers. Thus set theory served as a unifying principle of mathematics.

It has been noted, however, that some of the most commonly utilized arguments in set theory, which are at the same time the most useful in mathematics and belong almost to the basic framework of formal logic itself, resemble very much those which give rise to paradoxes. This fact has caused many critical mathematicians to question the very nature of mathematical reasoning. Thus a new field, **foundations of mathematics**, came into being at the beginning of this century. This field was divided at its inception into different doctrines according to the views of its initiators: **logicism** by B. Russell, **intuitionism** by L. E. J. Brouwer, and **formalism** by D. Hilbert. In set theory, which was the origin of this controversy, it was pointed out that the “definition” of set as given by G. Cantor was too naive, and axiomatic treatments of this theory were proposed (\rightarrow 33 Axiomatic Set Theory).

B. Logicism

Russell asserted that mathematics is a branch of logic and that paradoxes come from neglecting the “types” of concepts. According to his opinion, mathematics deals formally with structures independently of their concrete meanings. Science of this character has been called logic from antiquity. According to him, logic is the youth of mathematics, and mathematics is the manhood of logic. To construct mathematics from this standpoint, asserted Russell, ordinary language is lengthy and inaccurate, and some proper system of symbols should be used instead. Thus he tried to reconstruct mathematics using †symbolic logic.

Attempts to reorganize mathematics using logical symbols had formerly been made by G. Leibniz, who wrote *Dissertatio de arte combinatoria* in 1666, as well as by A. de Morgan, G. Boole, C. S. Peirce, E. Schröder, G. Frege, G. Peano, and others. Symbols used by the last two authors resemble those of today. Russell studied these works and published his own theory in a monumental joint work with A. N. Whitehead: *Principia mathematica* (3 vols., 1st ed. 1910–1913, 2nd ed. 1925–1927), in which the theories of natural numbers and real numbers as well as analytic geometry are developed from the fundamental laws of logic.

If this work had been completely successful, it could have eliminated any possibility of the intrusion of paradoxes into mathematics. However, the authors were forced to postulate an “unsatisfactory” axiom in order to construct mathematics. They introduced the notion of †type as follows: An object M defined as the set of all objects of a certain type belongs to a higher type than the types of the elements of M . This serves to eliminate certain paradoxes but brings about inconveniences such as the following. Suppose that we are trying to construct the theory of real numbers from that of rational numbers. Each real number can then be considered a †predicate about rational numbers. If this predicate contains only †quantifiers relating to variables running over all rational numbers, then the corresponding real number is said to be **predicative**, otherwise **impredicative**. According to Russell, the latter should have a higher type than the former, which makes the theory of real numbers exceedingly complicated. To avoid this difficulty, Russell proposed the **axiom of reducibility**, which says that every predicate can be replaced by a predicative one. With this rather artificial axiom Russell himself expressed dissatisfaction. Russell also postulated the †axiom of infinity and the †axiom of choice, which are also problematic. After examining the philosophical background of the book, H. Weyl wrote about *Principia mathematica*, “Mathematics is no more based on logic than the utopia built by the logician.” Nevertheless, logic as formulated in this book, as well as the theory of types as developed by F. P. Ramsay in the school of Russell and Whitehead, is still an important subject of mathematical logic.

C. Intuitionism

The intuitionist claims that mathematical objects or truths do not exist independently from mathematically thinking spirit or intuition, and that these objects or truths should be directly seized by mental or intuitional activ-

ity. The philosophical standpoints of mathematicians such as L. Kronecker and H. Poincaré in the 19th century or E. Borel, H. Lebesgue, and N. N. Luzin at the turn of this century can be assimilated to intuitionism, but those of the latter three are often said to belong to **semi-intuitionism** or to **French empiricism**. Brouwer took a narrower standpoint, strongly antagonistic to Hilbert’s formalism. Today the word “intuitionism” is generally interpreted in Brouwer’s sense.

Brouwer sharply criticized the usual way of reasoning in mathematics and claimed that indiscriminate use of the **law of excluded middle** (or **tertium non datur**) $P \vee \neg P$ cannot be permitted. According to him, the proposition “Either there exists a natural number with a given property P , or else no such number exists” is to be regarded as proved only when an actual construction of a natural number with the property P is given or when the absurdity of the existence of such a natural number can be constructively proved. When neither of these two results can be shown, then one can say nothing about the truth of the above proposition. Thus the usual method of proof, known as the method of **reductio ad absurdum**, i.e., of proving a proposition P by proving its double negation $\neg \neg P$, is not generally considered valid. It is a difficult but important problem of mathematical logic to determine which parts of usual mathematics can be reconstructed intuitionistically, though it does not seem easy to reconstruct any part of mathematics elegantly from this standpoint.

D. Formalism

To eliminate paradoxes, Hilbert tried to apply his axiomatic method. From Hilbert’s standpoint, any part of mathematics is a deductive system based on its axioms. In the deductive development, however, “logic,” including set theory and elementary number theory, is used. Paradoxes appear already in such logic. Hilbert’s idea was to axiomatize such logic and to prove its consistency. Thus one must first formalize the most elementary part of mathematics, including logic proper.

Hilbert proved the consistency of Euclidean geometry by assuming the consistency of the theory of real numbers. This is an example of a **relative consistency** proof, which reduces the consistency proof of one system to that of another. Such a proof can be meaningful only when the latter system can somehow be regarded as based on sounder ground than the former. To carry out the **consistency proof** of logic proper and set theory, one must reduce it to that of another system with sounder

ground. For this purpose, Hilbert initiated **metamathematics** and the **finitary standpoint**.

The finitary standpoint recognizes as its foundation only those facts that can be expressed in a finite number of symbols and only those operations that can be actually executed in a finite number of steps. Essentially, it does not differ from the standpoint of intuitionism. The methods based on this standpoint are also called **constructive methods**.

Metamathematics is also called **proof theory**. Its subject of research is mathematical proof itself. Hilbert was the first to insist on its importance. The theory is indispensable for consistency proofs of mathematical systems, but it may also be used for other purposes. In fact, the same idea can be seen in the †duality principle of projective geometry, which dates from long before Hilbert's proclamation of formalism. This is not a theorem of projective geometry deduced from its axioms; rather, it is a proposition about the theorems in projective geometry, based on the type of axioms and proofs in this subject.

According to Hilbert's method, one must develop proof theory from the finitary standpoint with the aim of proving the consistency of axiomatized mathematics. For this purpose, one must formalize the mathematical theory in question by means of symbolic logic. A theory thus formalized is called a **formal system**.

E. Some Results of Formalist Theory

One of the most remarkable results hitherto obtained with Hilbert's method is the consistency proof of pure number theory by G. Gentzen [7]. This consistency proof covers the largest domain for which an explicit consistency proof has so far been obtained. However, the methods of formalist proof theory have proved to be most effective in studying the logical structure of mathematical theories and have led to various results on the consistency of formalized mathematical systems, on symbolic logic, and on axiomatic set theory. We give some examples.

(1) Gödel's Incompleteness Theorem. K. Gödel [6] showed that if a system obtained by formalizing the theory of natural numbers is consistent, then this system contains a †closed formula A such that neither A nor its negation $\neg A$ can be proved within the system. He originally proved this under the assumption that the system is ω -consistent. This is a stronger condition for the system than simple consistency, but J. B. Rosser [13] succeeded in replacing this by the latter. This result shows the incompleteness not only of the usual

theory of natural numbers but of any consistent theory (from the finitary standpoint) containing the theory of these numbers.

At the same time, Gödel also obtained the following important result: Let S be any consistent formal system containing the theory of natural numbers. Then it is impossible to prove the consistency of S by utilizing only arguments that can be formalized in S . This means that a consistency proof from the finitary standpoint of a formal system S inevitably necessitates some argument that cannot be formalized in S .

(2) Consistency Proofs for Pure Number Theory. Gentzen [7] called **pure number theory** the theory of natural numbers not depending on the free use of set theory (differing consequently from the usual theory of natural numbers based on †Peano axioms; → 294 Numbers) and proved its consistency. W. Ackermann [14] proved the consistency of a similar theory admitting the use of Hilbert's † ε -symbol. G. Takeuti [15] showed that Gentzen's result can be obtained as a corollary to his theorem extending †Gentzen's fundamental theorem on †predicate logic of the first order to a †theory of types of a certain kind.

According to the result of Gödel mentioned in (1) above, some reasoning outside pure number theory must be used to prove its consistency. In all consistency proofs of pure number theory mentioned above, †transfinite induction up to the first † ε -number ε_0 is used, but all the other reasoning used in these proofs can be presented in pure number theory. This shows that the legitimacy of transfinite induction up to ε_0 cannot be proved in this latter theory. A direct proof of this fact was given by Gentzen [13]. On the other hand, the legitimacy of transfinite induction up to an ordinal number $< \varepsilon_0$ can be proved within pure number theory.

Again, transfinite induction is not the only method by which to prove the consistency of pure number theory. Actually, Gödel [17] carried out the proof utilizing what he called computable functions of finite type on natural numbers and what we call primitive recursive functionals of finite type.

By restricting pure number theory further, one obtains weaker theories of natural numbers whose consistency can be proved with finitary methods without recourse to such methods as transfinite induction up to ε_0 . M. Presburger [18] proved the consistency of a theory in which only the addition of numbers is considered an operation. Ackermann [19], J. von Neumann [20], J. Herbrand [21], and K. Ono [22] proved the consistency of theories in which some restrictions are placed

on the use of the axiom of \dagger mathematical induction.

On the other hand, K. Schütte [23] gave a consistency proof for number theory including what he called “infinite induction” from a stronger standpoint than Hilbert’s finitary one; he attempted to find a basis that makes such a proof possible.

(3) The Consistency of Analysis. No definitive result has yet been obtained from the standpoint of formalism, though many attempts are being made, among which a recent one by C. Spector [24] should be mentioned.

(4) Axiomatic Set Theory. There are different kinds of axiom systems (\rightarrow 33 Axiomatic Set Theory). To give a consistency proof for any of these systems is considered a very difficult problem today, but many interesting results are known concerning the relative consistency or independence of these axioms.

(5) The Skolem-Löwenheim Theorem. The metamathematical Skolem-Löwenheim theorem states: Given a consistent system of axioms stated in the first-order predicate logic whose cardinality is at most countable, there always exists an \dagger object domain consisting of countable objects satisfying all these axioms.

For example, axiomatic set theory is stated in predicate logic of the first order, and the cardinality of its axioms is countable. Thus there exists an object domain consisting of countable objects satisfying all these axioms, provided that they are consistent. Such a domain is called a **countable model** of axiomatic set theory. On the other hand, from the axioms of this theory one can prove that there exists a family of sets that is more than countable. This should also hold in a model of the theory, in which each object represents a set. This situation is known as the **Skolem paradox**.

This does not imply, however, the inconsistency of axiomatic set theory. In fact, the term “countable” is to be interpreted in its mathematical sense when one says “there exists a family of sets that is more than countable,” while it should be interpreted in its metamathematical sense when one speaks of a countable model of the theory. It is the confusion of these two different interpretations that leads to the “paradox.”

(6) Skolem’s Theorem on the Impossibility of Characterizing the System of Natural Numbers by Axioms. T. Skolem [27] proved that it is not possible to characterize the system of natural numbers by a countable system of axioms stated in the predicate logic of the first order. More precisely, given any consistent

countable system of axioms satisfied by the system of natural numbers, there always exists another \dagger linearly ordered system satisfying all these axioms and yet not isomorphic to the system of natural numbers as an ordered system.

Gödel’s incompleteness theorem and the Skolem paradox, as well as this result, seem to indicate that there is a certain limit to the effectiveness of the formalist method. On the other hand, nonstandard analysis has originated in this result.

References

- [1] S. C. Kleene, Introduction to metamathematics, Van Nostrand, 1952.
- [2] R. Dedekind, Was sind und was sollen die Zahlen? Vieweg, 1888.
- [3] R. Dedekind, Stetigkeit und irrationale Zahlen, Vieweg, 1872.
- [4] B. Russell, Introduction to mathematical philosophy, Allen & Unwin and Macmillan, 1919.
- [5] A. Heyting, Intuitionism, North-Holland, 1956.
- [6] K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, Monatsh. Math. Phys., 38 (1931), 173–198.
- [7] G. Gentzen, Die Widerspruchsfreiheit der reinen Zahlentheorie, Math. Ann., 112 (1936), 493–565.
- [8] D. Hilbert and P. Bernays, Grundlagen der Mathematik, Springer, I, second edition, 1968; II, 1970.
- [9] A. Tarski, Logic, semantics, metamathematics, Clarendon Press, 1956.
- [10] S. C. Kleene, Mathematical logic, Wiley, 1967.
- [11] J. R. Schoenfield, Mathematical logic, Addison-Wesley, 1967.
- [12] O. Becker, Grundlagen der Mathematik in geschichtlicher Entwicklung, Alber, 1954.
- [13] J. B. Rosser, Extensions of some theorems of Gödel and Church, J. Symbolic Logic, 1 (1936), 87–91.
- [14] W. Ackermann, Zur Widerspruchsfreiheit der Zahlentheorie, Math. Ann., 117 (1940), 162–194.
- [15] G. Takeuti, On the fundamental conjecture of GLC I, J. Math. Soc. Japan, 7 (1955), 249–275.
- [16] G. Gentzen, Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie, Math. Ann., 119 (1943), 140–161.
- [17] K. Gödel, Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes, Dialectica, 12 (1958), 280–287.

- [18] M. Presburger, Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt, C. R. du I Congrès des Math. des Pays Slaves, Warsaw (1929), 92–101, 395.
- [19] W. Ackermann, Begründung des “tertium non datur” mittels der Hilbertschen Theorie der Widerspruchsfreiheit, Math. Ann., 93 (1924), 1–36.
- [20] J. von Neumann, Zur Hilbertschen Beweistheorie, Math. Z., 26 (1927), 1–46.
- [21] J. Herbrand, Sur la non-contradiction de l'arithmétique, J. Reine Angew. Math., 166 (1931), 1–8.
- [22] K. Ono, Logische Untersuchungen über die Grundlagen der Mathematik, J. Fac. Sci. Univ. Tokyo, 3 (1938), 329–389.
- [23] K. Schütte, Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie, Math. Ann., 122 (1951), 369–389.
- [24] C. Spector, Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics, Recursive function theory, Amer. Math. Soc. Proc. Symp. Pure Math., 5 (1962), 1–27.
- [25] L. Löwenheim, Über Möglichkeiten in Relativkalkül, Math. Ann., 76 (1915), 447–470.
- [26] T. Skolem, Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre, 5 Congress der Skandinavischen Mathematiker, Helsingfors (1923), 217–232.
- [27] T. Skolem, Über die Nichtcharakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen, Fund. Math., 23 (1934), 150–161.

157 (VI.22) Four-Color Problem

A. Brief History

It is obvious that four colors are necessary to color some geographical maps on a sphere (or plane), but are four colors sufficient to color every map? This is the so-called four-color problem. The precise formulation of this problem will be given in Section B. The conjecture was made by Francis Guthrie and communicated through his brother Frederick Guthrie to A. De Morgan in 1852 [4]. A. Cayley called attention to the problem in 1878. It became famous after J. P. Heawood pointed out a mistake in the proof by A. B. Kempe (1879). Heawood also studied the problem of coloring maps on arbitrary surfaces (→ Section E).

Though there have been various approaches to the four-color problem, the main stream of investigation has concentrated on obtaining an unavoidable set consisting of reducible arrangements (→ Section D) in order to correct the mistake made by Kempe. G. D. Birkhoff [7] first discovered the simplest nontrivial reducible arrangement, nowadays called Birkhoff's diamond (Fig. 1). Using all reducible arrangements known up to that time, P. Franklin showed that every map with up to 25 countries must contain a reducible arrangement, so that such a map is four-colorable [8]. This limit has been gradually increased; in 1975 a reducible arrangement for 25 countries was found.

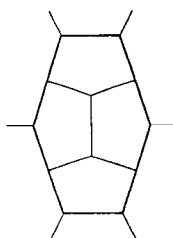


Fig. 1
Birkhoff's diamond.

H. Heesch invented the method of discharging (→ Section D), found criteria for reducibility, and finally conjectured the existence of an unavoidable set of reducible arrangements with several thousand elements, but this was too large to construct by hand. W. Haken and K. Appel with J. Koch worked with high-speed computers for a total of 1,200 CPU hours over a period of four and a half years and finally succeeded in constructing and checking an avoidable set of reducible arrangements with 1,834 elements, which proved the four-color problem affirmatively (1976; [9, 10]). For early investigations of the four-color problem → [5]; for results up to the 1930s → [3].

B. Precise Formulation of the Problem

To formulate the problem precisely, we must state the following two conditions: (i) Every country on a map is a †connected domain; a connected part of the sea is considered to be a country. (ii) Two countries sharing boundary lines must be colored differently. On the other hand, if two countries share only a finite number of points, then they may share the same color.

Actually we can modify the map so that no more than three countries meet at the same point. A map with this condition is called a **trivalent** or **cubic map**. In the study of the four-

color problem, we can restrict ourselves to trivalent maps without two-sided countries. In the remainder of this article, we shall assume that the maps considered have these properties.

In recent investigations, it has become customary to convert maps into their \dagger dual graphs, where each country is replaced by its capital lying inside, and for each adjacent pair A, B (Fig. 2), the boundary is replaced by a line connecting the representative points of A and B in such a way that it meets the boundary only once.

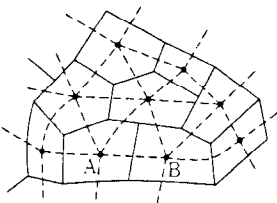


Fig. 2
Dual graph.

The formal extension of the four-color problem to higher dimensions is trivial, since we can easily construct arrangements needing arbitrary many colors for coloring. By converting to a dual graph, one can see that this is not surprising, since the planarity of a graph is a strong restriction, while the condition of representability of a linear graph in 3-dimensional space imposes no restriction.

C. Tait's Algorithm

Suppose that a planar trivalent map M is four-colored. Denote the four colors by 0, 1, 2, and 3. (In Fig. 3, we use A, B, C , and D instead of 0, 1, 2, and 3, respectively). Here we take the operation $a \oplus b \rightarrow c$. Represent the integers $a, b = 0, 1, 2, \text{ or } 3$ in the binary (2-adic) number system as 2-digit numbers $a_1 a_0$ and $b_1 b_0$, where a_i and b_i are 0 or 1. For each digit we take the operation $a_i \oplus b_i = c_i$ to be binary

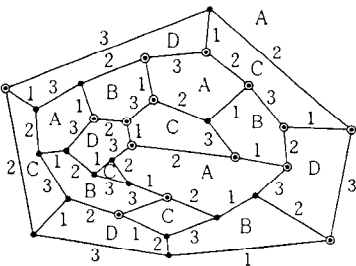


Fig. 3
An example of Tait's algorithm. At the vertices, ● stands for + and ○ for -. Within the domains, A, B, C, and D stand for four colors corresponding to 0, 1, 2, and 3, respectively.

addition without carrying, i.e., the operation is that of the logical (exclusive) "or." The number $c = c_1 c_0$ in the binary system represents the result of the operation. To each boundary line, we give the number corresponding to $a \oplus b$, where a and b are the colors for the two countries meeting at the boundary. Then we have the numbering 1, 2, and 3 for each edge, where 1, 2, and 3 are labeled once and only once for each triple of edges emanating from each vertex. Such numbering for edges is called **Tait coloring** for edges. Then we give the signature + or - to each vertex, according as the order of the edge numbering is counterclockwise or clockwise. Then the algebraic sum of the signatures along the boundary of each country in the map is always a multiple of 3. Conversely, if we give the signature + or - to each vertex in such a way that the algebraic sum of the signatures is always a multiple of 3 along the boundary of each country, then we get a four-coloring of the map by reversing the above procedure. This is called **Tait's algorithm**, which shows that the four-coloring of a planar map is an \dagger NP problem. As for five-coloring, it is known that an algorithm of polynomial complexity exists.

Tait's algorithm also shows that the four-color problem is equivalent to the following apparently elementary geometric proposition: "For any convex polyhedron, we can always cut near some of its vertices in such a way that the resulting polyhedron has only faces such that the number of sides is a multiple of 3." Many other equivalent statements of the four-color problem are known [1, 2].

D. Solution of the Four-Color Problem

For a given planar trivalent map, denote by V_n the number of the countries with n sides ($n \geq 3$). From Euler's theorem on polyhedra, we have immediately the relation

$$\sum_{n \geq 3} (6 - n) V_n = 12. \tag{1}$$

We easily see from this that every planar map must contain 3-, 4-, or 5-sided countries. A family F of the arrangements of countries with the property that every planar map must contain at least one arrangement belonging to F is called an **unavoidable set**. The family consisting of 3-, 4-, and 5-sided countries is the simplest example. In order to obtain an unavoidable set, Heesch invented the method of **discharging**. Let us assume the existence of a map M that contains no arrangement of a family F . We give a signed charge of $(6 - n)$ to every n -gon in M . Next we divide and move the charges so that the pluses and the minuses cancel out. If all positive charges disappear,

according to the assumption of the existence of the map M , then we have a contradiction of (1), so we can conclude that no such map M exists.

A 3-sided country can be ignored in four-coloring. As for a 4-sided country A , Kempe proved (by means of a definite modification procedure) that after four-coloring outside A , we also have a total four-coloring including A . An arrangement of a country or countries A is called **reducible** if, as in the above case, after four-coloring outside A , we get a total four-coloring including A by suitable modifications. If we have an unavoidable set consisting only of reducible arrangements, then by definition, every planar map will be four-colorable.

Kempe believed that he had proved the reducibility of a pentagon (5-sided country), but unfortunately, he missed a particular case. Even though the pentagon is not reducible, much effort has been given to finding other reducible arrangements. Many useful criteria for reducibility have also been studied. If we have enough reducible arrangements, then we can either eventually obtain an unavoidable set, which proves the four-color problem affirmatively, or find a “minimal counter-example” without reducible arrangements, which disproves it.

Haken obtained an unavoidable set consisting of arrangements without certain kinds of reduction obstructions. (He called these “geographically good arrangements.”) He firmly believed that, applying certain “probabilistic conjectures” to such arrangements, he would conquer the problem by considering arrangements of up to 14 countries. With repeated improvements on the discharging process and on the criteria for reducibility, he finally was able to conclude that his speculation was correct, as recounted in Section A. His investigation is not only an example of the use of high-speed computation in pure mathematics, but also one inviting reassessment of the meaning of mathematical proof.

E. Coloring Maps on Arbitrary Surfaces

On a \dagger torus, seven colors are sufficient to color any map and that many colors are necessary to color some maps. Heawood (1890) investigated the problem of coloring maps on closed surfaces with \dagger Euler characteristic $\chi < 2$, orientable or not. The least number of colors sufficient to color any map on a surface is called the **chromatic number** of the surface. Heawood proved that the chromatic number is less than or equal to

$$p = [(7 + \sqrt{49 - 24\chi})/2] \quad (\chi < 2), \tag{2}$$

where $[x]$ means the largest integer $a \leq x$. After this was proved in various special cases, J. W. T. Youngs and G. Ringel finally proved in 1968 that the number p equals the chromatic number except for a Klein bottle (a non-orientable surface with $\chi = 0$) [6]. Franklin proved in 1934 that the chromatic number for a Klein bottle is 6.

References

[1] O. Ore, The four-color problem, Academic Press, 1967.
[2] T. Saaty and P. Kainen, The four-color problem—assaults and conquest, McGraw-Hill, 1977.
[3] N. L. Biggs (ed.), Graph theory 1736–1936, Oxford Univ. Press, 1976.
[4] K. O. May, The origin of four-colour conjecture, *Isis*, 56 (1965), 346–348.
[5] W. W. R. Ball, Mathematical recreations and essays, Macmillan, 1892; revised twelfth edition, H. S. M. Coxeter (ed.), Toronto Univ. Press, 1958.
[6] G. Ringel, Map color theorem, Springer, 1974.
[7] G. D. Birkhoff, The reducibility of maps, *Amer. J. Math.*, 35 (1913), 115–128.
[8] P. Franklin, The four-color problem, *Amer. J. Math.*, 44 (1922), 225–236.
[9] K. Appel and W. Haken, Every planar map is four-colorable I. Discharging, *Illinois J. Math.*, 21 (1977), 429–489.
[10] K. Appel, W. Haken, and J. Koch, Every planar map is four-colorable II. Reducibility, *Illinois J. Math.*, 21 (1977), 490–567.

158 (XXI.24)
Fourier, Jean Baptiste Joseph

Jean Baptiste Joseph Fourier (March 21, 1768–May 16, 1830) was born in Auxerre, France, the son of a tailor, and was orphaned at the age of eight. In 1790, he was appointed professor at the Ecole Polytechnique. In 1798, Napoleon took him on his Egyptian campaign together with G. Monge. On his return to France, he was made governor of the department of Isère. With the downfall of Napoleon, he lost his position; however, he was later appointed to the French Academy of Science as a result of his research on the transmission of heat. In 1826 he was elected a member of the Académie Française.

His research on heat transmission was begun in 1800. In 1811, he presented a prize-

winning solution to a problem put forth by the Academy of Science. He solved the equation for heat transmission under various †boundary conditions. Fourier also stated (without rigorous proof) that an arbitrary function could be represented by †trigonometric series (→ 159 Fourier Series), a statement that gave rise to subsequent developments in analysis.

References

[1] J. B. J. Fourier, Oeuvres I, II, G. Darboux (ed.), Gauthier-Villars, 1888–1890.

159 (X.22) Fourier Series

Throughout this article, we assume that $f(x)$, $g(x)$, ... are real-valued functions and that integrals are always Lebesgue integrals.

A. Introduction

The set of functions

$$1/\sqrt{2\pi}, \quad \cos x/\sqrt{\pi}, \quad \sin x/\sqrt{\pi}, \dots, \\ \cos kx/\sqrt{\pi}, \quad \sin kx/\sqrt{\pi}, \dots,$$

which is called the **trigonometric system**, is an †orthonormal system in $(-\pi, \pi)$ (→ 317 Orthogonal Functions). Let $f(x)$ be an element of $L_1(-\pi, \pi)$ (i.e., †Lebesgue integrable in $(-\pi, \pi)$). We put

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos kt \, dt, \\ b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin kt \, dt, \quad k=0, 1, \dots, \quad (1)$$

and call a_k, b_k the **Fourier coefficients** of f . The formal series

$$\frac{1}{2}a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \quad (2)$$

is called the **Fourier series** of f and is often denoted by $\Xi(f)$. To indicate that a formal series $\Xi(f)$, as above, is the Fourier series of a function f , we write

$$f(x) \sim \frac{1}{2}a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx).$$

The sign \sim means that the numbers a_k, b_k are connected with f by the formula (1); it does not imply that the series is convergent, still less that it converges to f . Generally, **trigonometric series** are those of form (2), where the a_k, b_k are arbitrary real numbers. Since the trigono-

metric series have period 2π , we assume that the functions considered are extended for all real x by the condition of periodicity $f(x+2\pi)=f(x)$. The study of the properties of the series $\Xi(f)$ and the representation of f by $\Xi(f)$ are major objects of the theory of Fourier series. Since $e^{ix} = \cos x + i \sin x$, if we set $2c_k = a_k - ib_k$, $c_{-k} = \bar{c}_k$ ($k=0, 1, 2, \dots$), we have

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} \, dt, \quad k=0, \pm 1, \dots$$

Then $\Xi(f)$ is represented by the **complex form** $\sum_{k=-\infty}^{\infty} c_k e^{ikx}$, and $\{e^{ikx}\}$ ($k=0, \pm 1, \dots$) is an orthogonal system in $(-\pi, \pi)$. In this complex form, we take symmetric partial sums such as $\sum_{k=-n}^n c_k e^{ikx}$ ($n=1, 2, \dots$).

Consider the power series $\frac{1}{2}a_0 + \sum_{k=1}^{\infty} (a_k - ib_k)z^k$ on the unit circle $z = e^{ix}$ in the complex plane. Its real part is the trigonometric series (2), and the imaginary part (with vanishing constant term) is

$$\sum_{k=1}^n (a_k \sin kx - b_k \cos kx), \quad (3)$$

which is called the **conjugate series** (or **allied series**) of f and is denoted by $\tilde{\Xi}(f)$. In complex form, the conjugate series is $-i \sum_{k=-\infty}^{\infty} (\operatorname{sgn} k) c_k e^{ikx}$.

If f and g belong to $L_1(-\pi, \pi)$ and

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \quad g(x) \sim \sum_{k=-\infty}^{\infty} d_k e^{ikx},$$

then

$$\frac{1}{2\pi} \int_0^{2\pi} f(x-t)g(t) \, dt \sim \sum_{k=-\infty}^{\infty} c_k d_k e^{ikx}.$$

The function

$$f * g(x) = \frac{1}{2\pi} \int_0^{2\pi} f(x-t)g(t) \, dt$$

is called the **convolution** of f and g .

If f is †absolutely continuous, then the derivative $f'(x)$ satisfies

$$f'(x) \sim i \sum_{k=-\infty}^{\infty} k c_k e^{ikx} \\ = \sum_{k=1}^{\infty} k (b_k \cos kx - a_k \sin kx).$$

If $F(x)$ is an indefinite integral of f , then

$$F(x) - c_0 x \sim C + \sum_{k=-\infty}^{\infty} \frac{c_k}{ik} e^{ikx} \\ = C + \sum_{k=1}^{\infty} \frac{a_k \sin kx - b_k \cos kx}{k},$$

where C is a constant of integration and the symbol ' indicates that the term $k=0$ is omitted from the sum.

If $f \in L_1(-\pi, \pi)$, then the Fourier coefficients

c_n converge to 0 as $n \rightarrow \infty$ (**Riemann-Lebesgue theorem**). If f satisfies the † Lipschitz condition of order α ($0 < \alpha \leq 1$), then $c_n = O(n^{-\alpha})$, and if f is of † bounded variation, then $c_n = O(n^{-1})$.

When $f \in L_2(-\pi, \pi)$,

$$\begin{aligned} \frac{1}{\pi} \int_0^{2\pi} |f(x)|^2 dx &= 2 \sum_{k=-\infty}^{\infty} |c_k|^2 \\ &= \frac{1}{2} a_0^2 + \sum_{k=1}^{\infty} (a_k^2 + b_k^2), \end{aligned}$$

which is called the **Parseval identity**. If $\sum |c_k|^2 < \infty$, then there exists a function $f \in L_2(-\pi, \pi)$ which has the c_k as its Fourier coefficients.

This converse is implied by the † Riesz-Fischer theorem (\rightarrow Appendix A, Table 11.I).

B. Convergence Tests

The n th partial sums $s_n(x) = s_n(x; f)$ of the Fourier series $\mathfrak{S}(f)$ can be written in the form

$$s_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t) D_n(t) dt,$$

where

$$D_n(t) = \{\sin(n+1/2)t\} / 2 \sin(t/2).$$

The function $D_n(t)$ is called the **Dirichlet kernel**.

For a fixed point x we set $\varphi_x(t) = f(x+t) + f(x-t) - 2f(x)$; then

$$s_n(x) - f(x) = \frac{1}{\pi} \int_0^{\pi} \varphi_x(t) D_n(t) dt.$$

Hence if the integral on the right-hand side tends to zero as $n \rightarrow \infty$, $\lim_{n \rightarrow \infty} s_n(x) = f(x)$. If f vanishes in an interval $I = (a, b)$, then $\mathfrak{S}(f)$ converges uniformly in any interval $I' = (a + \varepsilon, b - \varepsilon)$ interior to I , and the sum of $\mathfrak{S}(f)$ is 0. This is called the **principle of localization**.

Here we give four convergence tests. (1) If f is of bounded variation, $\mathfrak{S}(f)$ converges at every point x to the value $\{f(x+0) + f(x-0)\}/2$. In addition, if f is continuous at every point of a closed interval I , $\mathfrak{S}(f)$ is uniformly convergent in I (**Jordan's test**). As a special case of this test, bounded functions having a finite number of maxima and minima and no more than a finite number of points of discontinuity have convergent Fourier series (**Dirichlet's test**). (2) If the integral $\int_0^{\pi} |\varphi_x(t)|/t dt$ is finite, then $\mathfrak{S}(f)$ converges at x to $f(x)$ (**Dini's test**). (3) If

$$\int_0^h |\varphi_x(t)| dt = o(h)$$

and

$$\lim_{\eta \rightarrow 0} \int_{\eta}^{\pi} \frac{|\varphi_x(t) - \varphi_x(t+\eta)|}{t} dt = 0,$$

then $\mathfrak{S}(f)$ converges at x to $f(x)$ (**Lebesgue's test**). Jordan's and Dini's tests are mutually independent, and both are included in Lebesgue's test, which, although not as convenient in certain cases, is quite powerful. (4) If $f(x)$ is continuous in (a, b) and its modulus of continuity satisfies the condition $\omega(\delta) \cdot \log(1/\delta) \rightarrow 0$ as $\delta \rightarrow 0$ in this interval, then $\mathfrak{S}(f)$ converges uniformly in $(a + \varepsilon, b - \varepsilon)$ (**Dini-Lipschitz test**).

C. Summability

Let $s_n(x)$ be the n th partial sum of the Fourier series $\mathfrak{S}(f)$, and $\sigma_n(x) = \sigma_n(x; f)$ be the first arithmetic mean ((C, 1)-mean) of $s_n(x)$ (i.e., $\sigma_n(x) = (s_0(x) + s_1(x) + \dots + s_n(x))/(n+1)$). Then we have

$$\sigma_n(x) - f(x) = \frac{1}{\pi} \int_0^{\pi} \varphi_x(t) K_n(t) dt,$$

where

$$K_n(t) = \frac{1}{2(n+1)} \left(\frac{\sin((n+1)t/2)}{\sin(t/2)} \right)^2.$$

The expression $K_n(t)$ is called the **Fejér kernel**, and the $\sigma_n(x)$ are often called **Fejér means**. If the right and left limits $f(x \pm 0)$ exist, $\mathfrak{S}(f)$ is † (C, 1)-summable at the point x to the value $(f(x+0) + f(x-0))/2$. If f is continuous at every point of a closed interval I , $\mathfrak{S}(f)$ is uniformly (C, 1)-summable in I (**Fejér theorem**, 1904). As we explain in Section H, there exist continuous functions whose Fourier series are divergent at some points. Thus the summability of $\mathfrak{S}(f)$ is more important than its convergence. Fejér's theorem remains true if we replace (C, 1)-summability by (C, α)-summability ($\alpha > 0$). More generally, if $f \in L_1(-\pi, \pi)$, then $\mathfrak{S}(f)$ is (C, α)-summable for $\alpha > 0$ to the value $f(x)$ almost everywhere (H. Lebesgue). Since (C, α)-summability ($\alpha > 0$) implies † summability by Abel's method, the result of Fejér's theorem is valid for † A-summability. However, the direct study of A-summability is also important. Let $f(r, x)$ be Abel's mean of $\mathfrak{S}(f)$; that is,

$$\begin{aligned} f(r, x) &= \frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) r^k \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t) P(r, t) dt, \end{aligned}$$

where $P(r, t) = (1 - r^2)/2(1 - 2r \cos t + r^2)$, $0 \leq r < 1$. We call $P(r, t)$ the **Poisson kernel**. The function $f(r, x)$ is † harmonic inside the unit circle and tends to $f(x)$ as $r \rightarrow 1$ almost everywhere. Hence $f(r, x)$ gives the solution of the † Dirichlet problem for the case of the unit circle.

D. The Gibbs Phenomenon

Let $f(x)$ be of bounded variation and not continuous at 0: $f(0)=0$, $f(+0)=l>0$, $f(-0)=-l$. Then the partial sum $s_n(x)$ converges to $f(x)$ in the neighborhood of 0, but not uniformly. Moreover,

$$\lim_{n \rightarrow \infty} s_n\left(\frac{\pi}{n}\right) = lG,$$

$$G = \frac{2}{\pi} \int_0^\pi \frac{\sin t}{t} dt = 1.1789 \dots$$

Hence as x tends to 0 from above and n tends to ∞ , the values $y = s_n(x)$ accumulate in the interval $[0, lG]$, while $s_n(-x) = -s_n(x)$ in the neighborhood of 0. This phenomenon is called the **Gibbs phenomenon**. If f is of bounded variation, then $\mathfrak{S}(f)$ exhibits the Gibbs phenomenon at every point of simple discontinuity of f . However, the $(C, 1)$ -means of $\mathfrak{S}(f)$ do not exhibit this phenomenon.

E. Conjugate Functions

For any integrable $f \in L_1(-\pi, \pi)$, the integral

$$\tilde{f}(x) =$$

$$\lim_{h \rightarrow 0} -\frac{1}{\pi} \int_h^\pi \left((f(x+t) - f(x-t)) / 2 \tan \frac{t}{2} \right) dt$$

exists almost everywhere. The function $\tilde{f}(x)$ is called the **conjugate function** of $f(x)$. The conjugate series $\mathfrak{S}(f)$ is (C, α) -summable ($\alpha > 0$) to the value $\tilde{f}(x)$ at almost every point, and a fortiori summable by Abel's method. Even if $f \in L_1(-\pi, \pi)$, \tilde{f} does not always belong to the class $L_1(-\pi, \pi)$. For example, $\sum_{n=2}^\infty \cos nx / \log n$ is the Fourier series of a function $f \in L_1(-\pi, \pi)$, but its conjugate series $\sum_{n=2}^\infty \sin nx / \log n$ is not the Fourier series of a function in $L_1(-\pi, \pi)$. However, if both f and \tilde{f} are integrable, $\mathfrak{S}(f) = \mathfrak{S}(\tilde{f})$. If $f \in L_p$ ($p > 1$), then $\tilde{f} \in L_p$ and $\|\tilde{f}\|_p \leq A_p \|f\|_p$; also, $\mathfrak{S}(f) = \mathfrak{S}(\tilde{f})$. If $|f| \log^+ |f|$ is integrable (such a function is said to belong to the **Zygmund class**), then \tilde{f} is integrable and $\mathfrak{S}(f) = \mathfrak{S}(\tilde{f})$. Moreover, in this case there exist constants A and B such that

$$\int_0^{2\pi} |\tilde{f}| dx \leq A \int_0^{2\pi} |f| \log^+ |f| dx + B.$$

If f is merely integrable, so is $|\tilde{f}|^p$ for any $0 < p < 1$, and $\|\tilde{f}\|_p \leq B_p \|f\|$ ($0 < p < 1$). If $f \in \text{Lip } \alpha$ ($0 < \alpha < 1$), then $\tilde{f} \in \text{Lip } \alpha$, but the theorem fails for $\alpha = 0$ and $\alpha = 1$. The conjugate function is important for convergence of partial sums of Fourier series.

F. Mean Convergence

The theorems on conjugate functions enable us to obtain some results for the mean convergence of the partial sums s_n of $\mathfrak{S}(f)$. If $f \in L_p$ ($p > 1$), then $\|f - s_n\|_p \rightarrow 0$; if $f \in L_1$, then $\|f - s_n\|_p \rightarrow 0$, $\|\tilde{f} - \tilde{s}_n\|_p \rightarrow 0$ for every $0 < p < 1$. Also, if $|f| \log^+ |f| \in L_1$, then $\|f - s_n\| \rightarrow 0$, $\|\tilde{f} - \tilde{s}_n\| \rightarrow 0$. As a corollary of this result, we obtain the following theorem, which is a generalization of the Parseval identity: If the Fourier coefficients of functions $f \in L_p$ and $g \in L_q$ ($1/p + 1/q = 1$) are a_n, b_n and a'_n, b'_n , respectively, we have the Parseval formula

$$\frac{1}{\pi} \int_0^{2\pi} fg dx = \frac{1}{2} a_0 a'_0 + \sum_{n=1}^\infty (a_n a'_n + b_n b'_n),$$

where this series is convergent.

G. Analytic Functions of the Class H_p

Let $p > 0$. A complex function $\varphi(z)$ holomorphic for $|z| < 1$ is said to belong to the class H_p (**Hardy class**) if there exists a constant M such that

$$\lim_{r \rightarrow 1} \frac{1}{2\pi} \int_0^{2\pi} |\varphi(re^{i\theta})|^p d\theta \leq M.$$

When $\varphi(z) \in H_p$, the nontangential limit $\varphi(e^{i\theta}) = \lim_{z \rightarrow e^{i\theta}} \varphi(z)$ exists for almost all θ . We write this as $\varphi(e^{i\theta}) = f(\theta) + i\tilde{f}(\theta)$, where $f(\theta), \tilde{f}(\theta)$ belong to the class L_p . Also, $\tilde{f}(\theta)$ coincides with the conjugate function of $f(\theta)$ for $p \geq 1$. For $1 < p < \infty$, H_p is isomorphic to L_p , but for $p = 1$ and $p = \infty$, H_p and L_p are different classes. Using the theory of functions of H_p , we can discuss some properties of Fourier series. If $\varphi(e^{i\theta}) = f(\theta) + i\tilde{f}(\theta)$ is of bounded variation, then $\varphi(e^{i\theta})$ is absolutely continuous and its Fourier series converges absolutely. We set

$$g(\theta) = \left(\int_0^1 (1-r) |\varphi'(re^{i\theta})|^2 dr \right)^{1/2},$$

$$g^*(\theta) =$$

$$\left(\int_0^1 (1-r) dr \int_0^{2\pi} |\varphi'(re^{i(\theta-t)})|^2 P(r, t) dt \right)^{1/2},$$

where $P(r, t)$ is the Poisson kernel. Then $g(\theta) \leq 2g^*(\theta)$, and there exist constants A_p, B, C , and A_μ such that

$$\int_0^{2\pi} |g(\theta)|^p d\theta \leq A_p \int_0^{2\pi} |\varphi(e^{i\theta})|^p d\theta, \quad p > 0,$$

$$\int_0^{2\pi} |g^*(\theta)|^p d\theta \leq A_p \int_0^{2\pi} |\varphi(e^{i\theta})|^p d\theta, \quad p > 1,$$

$$\int_0^{2\pi} |g^*(\theta)| d\theta \leq B \int_0^{2\pi} |\varphi| \log^+ |\varphi| d\theta + C,$$

$$\int_0^{2\pi} |g^*(\theta)|^\mu d\theta \leq A_\mu \left(\int_0^{2\pi} |\varphi(e^{i\theta})| d\theta \right)^\mu, \\ 0 < \mu < 1.$$

Denote by $s_n(\theta)$ and $\sigma_n(\theta)$, respectively, the partial sums and arithmetic means of the Fourier series of $\varphi(e^{i\theta})$, and set

$$\gamma(\theta) = \left(\sum_{n=1}^{\infty} \frac{|s_n(\theta) - \sigma_n(\theta)|^2}{n} \right)^{1/2}$$

Then $0 \neq A_1 \leq g^*(\theta)/\gamma(\theta) \leq A_2 \neq \infty$. From these relations, we can prove that if the indices n_k satisfy the conditions $\beta > n_{k+1}/n_k > \alpha > 1$, $s_{n_k}(\theta)$ converges almost everywhere to $\varphi(e^{i\theta})$ for $\varphi(e^{i\theta}) \in H_p$ ($1 \leq p$). If we set $\Delta_k(\theta) = \sum_{v=n_{k-1}+1}^{n_k} c_v e^{iv\theta}$, $\delta(\theta) = (\sum_{k=0}^{\infty} |\Delta_k(\theta)|^2)^{1/2}$, where $\varphi(e^{i\theta}) \sim \sum_{v=0}^{\infty} c_v e^{iv\theta}$, then $\|\delta(\theta)\|_p \leq A_p \|\varphi\|_p$ ($p > 1$). If $\varphi(z) \in H_p$ ($0 < p < 1$), then $\sum c_n e^{in\theta}$ is $(C, p^{-1} - 1)$ -summable to $\varphi(e^{i\theta})$ almost everywhere. These functions and relations were introduced mainly by J. E. Littlewood and R. E. A. C. Paley and were later generalized by A. Zygmund. There are more precise results by E. Stein [7], G. Sunouchi [8], S. Yano [9], and others.

H. Almost Everywhere Convergence and Divergence

P. du Bois Reymond (1876) first showed that there exists a continuous function whose Fourier series diverges at a point, but the problem of whether Fourier series of continuous functions converge almost everywhere (the so-called **du Bois Reymond problem**) remained unsolved for many years. At last in 1966, L. Carleson [10] proved that the Fourier series of a function belonging to L_2 converges almost everywhere; hence the du Bois Reymond problem was solved affirmatively. Using Carleson's method, R. A. Hunt [12] proved that

$$\int_0^{2\pi} \left(\sup_n |s_n(x)| \right)^p dx \leq A_p \int_0^{2\pi} |f(x)|^p dx, \\ 1 < p < \infty,$$

which implies that the Fourier series of $f \in L_p$ ($1 < p < \infty$) converges almost everywhere. Hunt also proved that

$$\int_0^{2\pi} \left(\sup_n |s_n(x)| \right) dx \\ \leq A \int_0^{2\pi} |f(x)| (\log^+ |f(x)|)^2 dx + A.$$

Moreover, P. Sjölín proved that if

$$\int_0^{2\pi} |f| \cdot \log^+ |f| \cdot \log^+ \log^+ |f| dx < \infty,$$

then the Fourier series of $f(x)$ converges almost everywhere.

On the other hand, A. N. Kolmogorov gave an integrable function with Fourier series diverging everywhere (more precisely, $\limsup_{n \rightarrow \infty} |s_{2^n}(x)| = \infty$ almost everywhere). Using this example, J. Marcinkiewicz showed that there is an $f \in L_1$ such that $s_n(x; f)$ oscillates boundedly almost everywhere. Moreover, there exists an integrable function with integrable conjugate and almost everywhere diverging Fourier series [3]. For any given null set E , we can construct a continuous function whose Fourier series diverges on every $x \in E$.

I. Absolute Convergence

The convergence of the series (1) $\sum(|a_n| + |b_n|)$ implies the absolute convergence of the trigonometric series (2) $a_0/2 + \sum(a_n \cos nx + b_n \sin nx)$. Conversely, if the series (2) converges absolutely in a set of positive measure, the series (1) converges (**Denjoy-Luzin theorem**). For the absolute convergence of Fourier series, we have the following tests: If $f \in \text{Lip } \alpha$ ($\alpha > 1/2$), then $\mathfrak{S}(f)$ converges absolutely, but for $\alpha = 1/2$, this is no longer true. If $f(x)$ is of bounded variation and belongs to $\text{Lip } \alpha$ ($\alpha > 0$), $\mathfrak{S}(f)$ converges absolutely.

Suppose that the Fourier series of a function $f(x)$ is absolutely convergent and the value of $f(x)$ belongs to an interval (a, b) . If $\varphi(z)$ is a function of a complex variable holomorphic at every point of the interval (a, b) , the Fourier series of $\varphi\{f(x)\}$ converges absolutely (**Wiener-Lévy theorem**). As a corollary we obtain that if $\mathfrak{S}(f)$ converges absolutely and $f(x) \neq 0$, then $\mathfrak{S}(1/f)$ converges absolutely. The converse of the Wiener-Lévy theorem was proved by Y. Katznelson [6]. For a given $\varphi(x)$ defined in $[-1, 1]$, if the Fourier series of $\varphi\{f(x)\}$ converges absolutely for every $f(x)$ with absolutely convergent Fourier series ($|f(x)| \leq 1$), then $\varphi(z)$ is holomorphic at every point of the interval $[-1, 1]$.

Many problems concerning this topic still remain unsolved. In particular, the determination of the structure of the functions with absolutely convergent Fourier series has not been completed.

J. Sets of Uniqueness

If $a_n \cos nx + b_n \sin nx$ converges to 0 on a set of positive measure, then $a_n, b_n \rightarrow 0$ (**Cantor-Lebesgue theorem**). A point set $E \subset (0, 2\pi)$ is called a **set of uniqueness** (or **U-set**) if every trigonometric series converging to 0 outside E vanishes identically. A set that is not a U-set is

called a **set of multiplicity** (or **M-set**). G. Cantor showed that every finite set is a U-set, and W. H. Young showed that every denumerable set is a U-set. It is clear that any set E of positive measure is an M-set, but D. E. Men'shov showed that there are \dagger perfect M-sets of measure 0. Moreover, N. K. Bari showed that there exist perfect sets of type U. However, the structure problem of sets of uniqueness has not yet been solved completely.

A set E is said to be of type H (or an H-set) if there exists a sequence of positive integers $n_1 < n_2 < \dots$ and an interval I such that for each $x \in E$, no point of $\{n_k x\}_{k=1}^\infty$ is in $I \pmod{2\pi}$. H-sets are sets of uniqueness, a fact given by A. Rajchman. I. I. Pyatetskii-Shapiro generalized H-sets to $H^{(m)}$ -sets [3].

Lacunary trigonometric series are series in which very few terms differ from zero. Such series can be written in the form

$$\sum_{k=1}^{\infty} (a_k \cos n_k x + b_k \sin n_k x) = \sum_{k=1}^{\infty} A_{n_k}(x).$$

S. Sidon established some of the characteristic properties of such series; he generalized them further and obtained the notion of Sidon sets (\rightarrow 192 Harmonic Analysis). We often define a lacunary series more specifically as a series for which the n_k satisfy Hadamard's gaps; that is, $n_{k+1}/n_k > q > 1$. Then if $\sum_{k=1}^{\infty} (a_k^2 + b_k^2)$ is finite, the series $\sum_{k=1}^{\infty} A_{n_k}(x)$ converges almost everywhere. Conversely, if $\sum_{k=1}^{\infty} A_{n_k}(x)$ is convergent in a set of positive measure, then $\sum_{k=1}^{\infty} (a_k^2 + b_k^2)$ converges. This theorem is related to the Rademacher series and random Fourier series [4].

K. Multiple Fourier Series

Routine extensions to multiple Fourier series from the case of a single variable are easy, but significant results are difficult to obtain. Recently, however, there have been several important contributions in this field.

References

- [1] A. Zygmund, *Trigonometrical series*, Warsaw, 1935.
- [2] A. Zygmund, *Trigonometric series I, II*, Cambridge Univ. Press, second edition, 1959.
- [3] G. H. Hardy and W. W. Rogosinski, *Fourier series*, Cambridge Univ. Press, 1950.
- [4] N. K. Bari, *A treatise on trigonometric series*, Pergamon, 1964. (Original in Russian, 1958.)
- [5] J.-P. Kahane and R. Salem, *Ensembles parfaits et séries trigonométriques*, *Actualités Sci. Ind.*, Hermann, 1963.
- [6] Y. Katznelson, *Sur les fonctions opérant*

sur l'algèbre des séries de Fourier absolument convergentes, *C. R. Acad. Sci. Paris*, 247 (1958), 404–406.

[7] E. M. Stein, *A maximal function with applications to Fourier series*, *Ann. Math.*, (2) 68 (1958), 584–603.

[8] G. Sunouchi, *Theorems on power series of the class H^p* , *Tôhoku Math. J.*, (2) 8 (1956), 125–146.

[9] S. Yano, *On a lemma of Marcinkiewicz and its applications to Fourier series*, *Tôhoku Math. J.*, (2) 11 (1959), 191–215.

[10] L. Carleson, *On convergence and growth of partial sums of Fourier series*, *Acta Math.*, 116 (1966), 135–157.

[11] J.-P. Kahane and Y. Katznelson, *Sur les ensembles de divergence des séries trigonométriques*, *Studia Math.*, 26 (1966), 305–306.

[12] R. A. Hunt, *On the convergence of Fourier series, Orthogonal expansions and their continuous analogues*, *Proc. Conf. Southern Illinois Univ.*, 1967, Southern Illinois Univ. Press, 1968, 235–255.

[13] P. Sjölin, *An inequality of Paley and convergence a.e. of Walsh-Fourier series*, *Ark. Mat.*, 7 (1968), 551–570.

[14] R. Salem, *Algebraic numbers and Fourier analysis*, Heath, 1963.

[15] J.-P. Kahane, *Some random series of functions*, Heath, 1968.

160 (X.23) Fourier Transform

A. Fourier Integrals

In this article we assume that $f(x)$ is a complex-valued function defined on $\mathbf{R} = (-\infty, \infty)$ and \dagger (Lebesgue) integrable on any finite interval. If the integral

$$\int_{-\infty}^{\infty} f(x) e^{-ixt} dx = \lim_{\substack{A \rightarrow -\infty \\ B \rightarrow \infty}} \int_A^B f(x) e^{-ixt} dx$$

exists, it is called the **trigonometric integral** or **Fourier integral**. We have a general result: If $f(x) \in L_1(-\infty, \infty)$, $K(x)$ is bounded on $(-\infty, \infty)$, and $\int_0^T K(x) dx = o(T)$ ($T \rightarrow \pm\infty$), then $\int_{-\infty}^{\infty} f(x) K(xt) dx$ exists and

$$\lim_{t \rightarrow \pm\infty} \int_{-\infty}^{\infty} f(x) K(xt) dx = 0.$$

In particular, it follows that if $f(x) \in L_1(-\infty, \infty)$, then $\int_{-\infty}^{\infty} f(x) e^{itx} dx$ exists and

$$\lim_{t \rightarrow \pm\infty} \int_{-\infty}^{\infty} f(x) e^{-itx} dx = 0$$

(Riemann-Lebesgue theorem).

B. Fourier's Integral Theorems

Suppose that $f(x)$ is of †bounded variation in an interval including x , or more generally satisfies the assumption for any one of the convergence tests for Fourier series (→ 159 Fourier Series B). Then **Fourier's single integral theorem**

$$\frac{1}{2}(f(x+0)+f(x-0)) = \lim_{A \rightarrow \infty} \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \frac{\sin A(t-x)}{t-x} dt \quad (1)$$

holds if one of the following three conditions is satisfied: (1) $f(x)/(1+|x|)$ belongs to $L_1(-\infty, \infty)$; (2) $f(x)/x$ tends to zero monotonically as $x \rightarrow \pm\infty$; (3) $f(x)/x = g(x)\sin(px+q)$, where $g(x)$ tends to zero monotonically as $x \rightarrow \pm\infty$ (S. Izumi, 1934). The right-hand side of (1) is called **Dirichlet's integral**.

Let $f(x)$ be of bounded variation in an interval including x (or satisfy some other convergence test for Fourier series). Then **Fourier's double integral theorem**

$$\frac{1}{2}(f(x+0)+f(x-0)) = \frac{1}{\pi} \lim_{T \rightarrow \infty} \int_0^T dt \int_{-\infty}^{\infty} f(u) \cos t(u-x) du \quad (2)$$

holds if one of the following three conditions is satisfied: (4) $f(x) \in L_1(-\infty, \infty)$; (5) $f(x)/(1+|x|) \in L_1(-\infty, \infty)$, and $f(x)$ tends to zero monotonically as $x \rightarrow \pm\infty$; (6) $f(x)/(1+|x|) \in L_1(-\infty, \infty)$ and $f(x) = g(x)\sin(px+q)$, where $g(x)$ tends to zero monotonically as $x \rightarrow \pm\infty$. If $f(x) \in L_1(-\infty, \infty)$, the formula

$$f(x) = \lim_{A \rightarrow \infty} \frac{1}{\pi A} \int_{-\infty}^{\infty} f(t) \frac{\sin^2 A(t-x)}{A(t-x)^2} dt$$

holds almost everywhere, and in particular at any x where $f(x)$ is continuous. More generally, the formula

$$f(x) = \lim_{A \rightarrow \infty} A \int_{-\infty}^{\infty} f(t) K(A(t-x)) dt = \lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} f\left(x - \frac{t}{A}\right) K(t) dt \quad (3)$$

holds at any point x where $f(x+0)$ and $f(x-0)$ exist, if $K(t) \in L_1(-\infty, \infty)$, $\int_{-\infty}^{\infty} K(t) dt = 1$, $|K(t)| < M$, $K(t) = o(t^{-1})$ as $|t| \rightarrow \infty$, and $f(x) \in L_1(-\infty, \infty)$, or if $K(t) \in L_1(-\infty, \infty)$, $\int_{-\infty}^{\infty} K(t) dt = 1$, and $f(x)$ is bounded. Similarly, the formula

$$\lim_{A \rightarrow 0} \int_0^{\infty} f\left(\frac{x}{A}\right) K(x) dx = \mathfrak{M}\{f\} \int_0^{\infty} K(x) dx$$

holds if $\mathfrak{M}\{f\} = \lim_{t \rightarrow \infty} T^{-1} \int_0^T f(t) dt$ exists, $K(x)$ is differentiable, $|x^2 K(x)| \leq C$ ($1 \leq x$), and

$x^{-1} \int_0^x |f(t)| dt \leq D$, where C, D are constants (**Wiener's formula**).

C. Fourier Transforms (→ Appendix A, Table 11.II)

Let $f(x) \in L_1(0, \infty)$. Then

$$F(t) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(u) \cos ut du$$

is called the **(Fourier) cosine transform** of $f(x)$. Under the same condition as for the validity of (2), the **inversion formula** of the cosine transform $f(x) = \sqrt{2/\pi} \int_0^{\infty} F(t) \cos xt dt$ holds where we suppose that $f(x) = \frac{1}{2}(f(x+0) + f(x-0))$. If we define $f(-x) = f(x)$, then this is equivalent to the formula (2). Analogously,

$$G(t) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(u) \sin ut du$$

is called the **(Fourier) sine transform** of $f(x)$. Under the same condition as for the validity of (2), we get the inversion formula $f(x) = \sqrt{2/\pi} \int_0^{\infty} G(t) \sin xt dt$. More generally, for any $f(x) \in L_1(-\infty, \infty)$,

$$F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ixt} dx$$

is called the **Fourier transform** of $f(x)$. Under the same condition as for the validity of (2), the **Fourier inversion formula**

$$f(x) = \frac{1}{\sqrt{2\pi}} \lim_{T \rightarrow \infty} \int_{-T}^T F(t) e^{ixt} dt$$

holds. The cosine transform and the sine transform coincide with the Fourier transform when $f(x) = f(-x)$ and $-f(x) = f(-x)$, respectively.

If for any $f(x)$, $F(t) \in L_q(-\infty, \infty)$ ($1 \leq q < \infty$) exists, for which

$$\int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{2\pi}} \int_{-T}^T f(x) e^{-ixt} dx - F(t) \right|^q dt \rightarrow 0 \quad (T \rightarrow \infty)$$

is valid (i.e., $(1/\sqrt{2\pi}) \int_{-T}^T f(t) e^{-ixt} dt$ converges to $F(x)$ as $T \rightarrow \infty$ in the mean of order q), then we say that $f(x)$ has the **Fourier transform** $F(t)$ in $L_q(-\infty, \infty)$. If $f(x) \in L_p(-\infty, \infty)$ ($1 < p \leq 2$), then $f(x)$ has the Fourier transform $F(t)$ in L_q ($1/p + 1/q = 1$), and $F(t)$ has the Fourier transform $f(-x)$ in L_p (E. C. Titchmarsh). Moreover,

$$\int_{-\infty}^{\infty} |F(t)|^q dx \leq \frac{1}{(2\pi)^{(q/2)-1}} \left(\int_{-\infty}^{\infty} |f(x)|^p dx \right)^{1/(p-1)}.$$

If $f(x), G(x) \in L_p(-\infty, \infty)$ ($1 < p \leq 2$) and their Fourier transforms in L_q are $F(t), g(t)$, respectively, then the **Parseval identity**

$$\int_{-\infty}^{\infty} F(t)G(t)dt = \int_{-\infty}^{\infty} f(x)g(-x)dx$$

holds. The Fourier inversion formula holds, in the sense that

$$\int_0^x F(u)du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) \frac{e^{-ixt} - 1}{-it} dt,$$

$$\int_0^x f(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(u) \frac{e^{ixu} - 1}{iu} du.$$

The theory of Fourier transforms is valid for cosine and sine transforms. Specifically for $p = 2$, if $f \in L_2(0, \infty)$, then $\sqrt{2/\pi} \int_0^T f(x) \cdot \cos xt dt$ converges in the mean in $L_2(0, \infty)$ to the cosine transform $F(t)$ as $T \rightarrow \infty$, and conversely, the cosine transform of $F(t)$ in L_2 is $f(x)$. The transforms $f(x), F(x)$ are connected by the formulas

$$\int_0^x F(u)du = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(t) \frac{\sin xt}{t} dt,$$

$$\int_0^x f(t)dt = \sqrt{\frac{2}{\pi}} \int_0^{\infty} F(u) \frac{\sin xu}{u} du,$$

$$\int_0^{\infty} |f(x)|^2 dx = \int_0^{\infty} |F(t)|^2 dt.$$

The theory of Fourier transforms was generalized as follows by G. N. Watson (*Proc. London Math. Soc.*, 35 (1933)). We suppose that $\chi(x)/x \in L_2(0, \infty)$ and

$$\int_0^{\infty} \frac{\chi(xu)\chi(yu)}{u^2} du = \min(x, y). \quad (4)$$

If $f(x) \in L_2(0, \infty)$, then there exists an $F(t) \in L_2(0, \infty)$ such that

$$\int_0^x F(t)dt = \int_0^{\infty} \frac{\chi(xu)f(u)}{u} du,$$

and the inversion formula

$$\int_0^x f(u)du = \int_0^{\infty} \frac{\chi(xt)F(t)}{t} dt$$

and the Parseval identity

$$\int_0^{\infty} |f(x)|^2 dx = \int_0^{\infty} |F(t)|^2 dt$$

hold. $F(t)$ is called the **Watson transform** of $f(x)$. For any $f \in L_2$, equality (4) is necessary for the existence of the Watson transform $F(t)$ for which the inversion formula holds. S. Bochner (1934) generalized this theory further to unitary transformations in L_2 (\rightarrow 192 Harmonic Analysis).

D. Conjugate Functions

Corresponding to Fourier's double integral theorem (2), the integral

$$\lim_{\lambda \rightarrow \infty} \frac{1}{\pi} \int_0^{\lambda} dt \int_{-\infty}^{\infty} \sin t(u-x)f(u)du$$

is called the **conjugate Fourier integral** of the integral in the right-hand side of (2). Formally, this is written $\lim_{x \rightarrow \infty} t^{-1} \int_0^{\infty} (1 - \cos \lambda t) t^{-1} (f(x+t) - f(x-t)) dt$. If $f(x)$ is a sufficiently regular function, the part involving $\cos \lambda t$ tends to 0 as $\lambda \rightarrow \infty$. Now let

$$g(x) = \frac{1}{\pi} \lim_{\substack{A \rightarrow \infty \\ \varepsilon \rightarrow 0}} \int_{\varepsilon}^A \frac{f(x+t) - f(x-t)}{t} dt.$$

For any $f \in L_1(0, \infty)$, the integral exists almost everywhere, and $g(x)$ is called the **conjugate function** or **Hilbert transform** of $f(x)$. If $f \in L_p$ ($p > 1$), then $g(x) \in L_p$ also and we have

$$f(x) = -\frac{1}{\pi} \lim_{\substack{A \rightarrow \infty \\ \varepsilon \rightarrow 0}} \int_{\varepsilon}^A \frac{g(x+t) - g(x-t)}{t} dt$$

and $\int_{-\infty}^{\infty} |g(x)|^p dx \leq M_p \int_{-\infty}^{\infty} |f(x)|^p dx$, where M_p is a constant depending only on p . In particular,

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |g(x)|^2 dx \quad \text{for } p=2.$$

E. Boundary Functions of Analytic Functions

Suppose that a complex-valued function $f(z)$ ($z = x + iy$) is holomorphic for $y > 0$, $f(x + iy)$ converges as $y \rightarrow 0$ for almost all x to $f(x)$ (which is called the **boundary function**), and $f(x) \in L_p(-\infty, \infty)$ ($p \geq 1$). Moreover, suppose that $f(z)$ is represented by Cauchy's integral formula or Poisson's integral of $f(x)$ on the real line. If either $f(x) \in L_p$ ($p \geq 1$) and has $F(t)$ as its Fourier transform or $f(x)$ is an L_p -Fourier transform of $F(t) \in L_q$ ($q \geq 1$), then a necessary and sufficient condition for the function $f(x)$ to be the boundary function of an analytic function is that $F(t)$ be 0 almost everywhere for $t > 0$ (N. Wiener, R. E. A. C. Paley, E. Hille, J. D. Tamarkin).

F. Generalized Fourier Integrals

Let $|f(x)|/(1 + |x|^k) \in L_1(-\infty, \infty)$ for a positive integer k and

$$L_k = L_k(t, x) = \begin{cases} \sum_{v=0}^{k-1} \frac{(-itx)^v}{v!}, & |x| \leq 1, \\ 0, & |x| > 1, \end{cases}$$

$$E_k(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \frac{e^{-itx} - L_k}{(-ix)^k} dx.$$

The function $E_k(t)$ or one that differs from $E_k(t)$ by a polynomial of degree at most k is called the k th transform of $f(x)$. We write formally $f(x) = \int_{-\infty}^{\infty} e^{ixt} d^k E_k(t)$. Actually, if we give an appropriate meaning to the integral, this formula itself is valid (H. Hahn, *Wiener Berichte*, 134 (1925); S. Izumi, *Tôhoku Science Rep.*, 23 (1935)). (For the theory and applications of k th transforms \rightarrow [2, ch. 6].)

G. Applications of Fourier Transforms

Suppose that $f(x) \in L_1(-\infty, \infty)$, $F(t)$ is its Fourier transform, and $f(x) = o(e^{-\theta(x)})$, where $\theta(x)$ is positive and increasing. If $\int_1^\infty \theta(x)x^{-2} dx = \infty$ and $F(t)$ vanishes identically in an interval, then $F(t) \equiv 0$ over $(-\infty, \infty)$. If $\int_1^\infty \theta(x)x^{-2} dx < \infty$, then there exists a function $f(x)$ such that $F(t)$ vanishes identically in an interval, but $F(t)$ does not vanish identically in $(-\infty, \infty)$ (Wiener, Paley, N. Levinson). These results are applicable to the theory of † quasi-analytic functions.

Let $f(x) \in L_1(-\infty, \infty)$. Then a necessary and sufficient condition for any function in $L_1(-\infty, \infty)$ to be approximated as closely as we wish by linear combinations of the translations $\sum_{k=1}^N a_k f(x+h_k)$ of $f(x)$ with respect to the L_1 -norm is that the Fourier transform of $f(x)$ does not vanish at any real number. When $f(x) \in L_2(-\infty, \infty)$, a necessary and sufficient condition that an arbitrary function in $L^2(-\infty, \infty)$ can be approximated as closely as we wish by $\sum_{k=1}^N a_k f(x+h_k)$ with respect to the L_2 -norm is that the zeros of the Fourier transform of $f(x)$ have measure zero (Wiener). This result was used by Wiener to prove the **generalized Tauberian theorem**: Suppose that $g_1(x) \in L_1(-\infty, \infty)$ and its Fourier transform never vanishes. Moreover, let $g_2(x) \in L_1(-\infty, \infty)$ and $p(x)$ be bounded over $(-\infty, \infty)$. Then $\lim_{x \rightarrow \infty} \int_{-\infty}^x g_1(x-t)p(t) dt = A \int_{-\infty}^{\infty} g_1(t) dt$ implies that $\lim_{x \rightarrow \infty} \int_{-\infty}^x g_2(x-t)p(t) dt = A \int_{-\infty}^{\infty} g_2(t) dt$. Another type of Wiener theorem is concerned with † Stieltjes integrals. Suppose that

$$\sum_{n=-\infty}^{\infty} \sup_{n \leq x < n+1} |g_1(x)| < \infty$$

(hence $g_1(x) \in L_1$) and that the Fourier transform of $g_1(x)$ never vanishes. Moreover, let

$$\sum_{n=-\infty}^{\infty} \sup_{n \leq x < n+1} |g_2(x)| < \infty$$

and let $\int_x^{x+1} |d\alpha(t)|$ be bounded. Then

$$\lim_{x \rightarrow \infty} \int_{-\infty}^x g_1(x-t) d\alpha(t) = A \int_{-\infty}^{\infty} g_1(t) dt$$

implies

$$\lim_{x \rightarrow \infty} \int_{-\infty}^x g_2(x-t) d\alpha(t) = A \int_{-\infty}^{\infty} g_2(t) dt.$$

From these general theorems, we can prove various † Tauberian theorems about the summation of series. Also, these results were applied to the proof of the † prime number theorem by S. Ikehara and E. Landau (Wiener [1]). In the general Tauberian theorem, the boundedness of $p(t)$ can be replaced by one-sided boundedness (H. R. Pitt, 1938). In fact, the first form of the theorem still holds if we replace the condition by the following one: g_1 and g_2 are continuous, $g_1(x) \geq 0$, the Fourier transform of $g_1(x)$ does not vanish, $g_2(x)$ satisfies the conditions of the second theorem, and $p(x) \geq C$. (Concerning Fourier transforms on the topological groups \rightarrow 192 Harmonic Analysis; and concerning Fourier transforms of the distributions \rightarrow 125 Distributions and Hyperfunctions).

H. Fourier Transforms of Distributions

The Fourier transform is defined in higher-dimension \mathbf{R}^n by

$$\mathcal{F}(f) = (\sqrt{2\pi})^{-n} \int e^{-ix\xi} f(x) dx, \quad f \in L_1,$$

$$x\xi = x_1\xi_1 + x_2\xi_2 + \dots + x_n\xi_n.$$

One denotes it by $\hat{f}(\xi)$. The inverse transform is defined by

$$\overline{\mathcal{F}}(g) = (\sqrt{2\pi})^{-n} \int e^{ix\xi} g(\xi) d\xi.$$

Differentiation under the integral sign gives

$$D^\alpha \hat{f}(\xi) = (\sqrt{2\pi})^{-n} \int e^{-ix\xi} (-ix)^\alpha f(x) dx,$$

$$D^\alpha = (\partial/\partial x_1)^{\alpha_1} \dots (\partial/\partial x_n)^{\alpha_n},$$

under the assumption $(1+|x|)^{|\alpha|} f(x) \in L_1$ ($|\alpha| = \alpha_1 + \dots + \alpha_n$). Roughly speaking, the decreasing order of $f(x)$ when $|x| \rightarrow \infty$ is reflected in the differentiability of $\hat{f}(\xi)$. In the same way,

$$(i\xi)^\alpha \hat{f}(\xi) = (\sqrt{2\pi})^{-n} \int e^{-ix\xi} D^\alpha f(x) dx,$$

which shows that the differentiability of $f(x)$ is reflected in the decreasing order of $\hat{f}(\xi)$.

The same statements can evidently be made for the relation between $g(\xi)$ and its inverse Fourier transform.

Let \mathcal{S} be the space of rapidly decreasing functions $\varphi(x)$, i.e., such that for all positive integers k and $\alpha \geq 0$, $(1+|x|)^k D^\alpha \varphi(x)$ remains bounded (\rightarrow 125 Distributions and Hyperfunctions). From the facts above, it follows that the

linear mapping $f(x) \mapsto \mathcal{F}(f)$ is a topological isomorphism from \mathcal{S}_x onto \mathcal{S}_ξ . Let \mathcal{S}' be the dual space of \mathcal{S} . Usual function spaces in classical analysis are contained in \mathcal{S}' , for instance, the L_p -space and that of functions increasing at most of polynomial order at infinity.

Let $T \in \mathcal{S}'_x$. Then the mapping $\varphi \in \mathcal{S}_\xi \mapsto T(\mathcal{F}\varphi)$ is continuous, and $\mathcal{F}T \in \mathcal{S}'_\xi$ is defined by $\mathcal{F}T(\varphi) = T(\mathcal{F}\varphi)$. \mathcal{F} is also a topological isomorphism from \mathcal{S}'_x onto \mathcal{S}'_ξ . Take as an example the distribution p.v. $1/x$. This is no longer a function; however, its Fourier transform can be calculated as follows:

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ A \rightarrow \infty}} (\sqrt{2\pi})^{-1} \int_{\varepsilon \leq |x| \leq A} e^{-ix\xi} \frac{dx}{x} = \frac{2i}{\sqrt{2\pi}} \\ \times \lim_{A \rightarrow \infty} \int_0^A \frac{\sin(x\xi)}{x} dx = \begin{cases} \sqrt{\frac{\pi}{2}} i, & \xi > 0, \\ -\sqrt{\frac{\pi}{2}} i, & \xi < 0, \end{cases}$$

where the limit is taken with respect to the topology of \mathcal{S}' .

Using the definition above, classical results can be extended to \mathcal{S}' almost automatically. For example, $\mathcal{F}(D^\alpha T) = (i\xi)^\alpha \mathcal{F}(T)$, $\mathcal{F}((-ix)^\alpha T) = D^\alpha \mathcal{F}(T)$. Moreover, if $T \in \mathcal{S}'$, then $\mathcal{F}T = \sqrt{2\pi}^{-n} T(e^{-ix\xi})$. In particular, $\mathcal{F}\delta = (2\pi)^{-n/2}$.

The **Plancherel theorem** says that if for $f(x) \in L_2$ one defines its Fourier transform by

$$\hat{f}(\xi) = \text{l.i.m.}_{A \rightarrow \infty} \sqrt{2\pi}^{-n} \int_{|x| \leq A} e^{-ix\xi} f(x) dx,$$

then the correspondence $f(x) \mapsto \hat{f}(\xi)$ is a unitary mapping from L_2 onto L_2 , i.e.,

$$\int |f(x)|^2 dx = \int |\hat{f}(\xi)|^2 d\xi.$$

This result can be extended. For any non-negative integer m , the element of H^m (\rightarrow 168 Function Spaces) is characterized by its Fourier transform: $f(x) \in H^m$ if and only if $(1 + |\xi|)^m \hat{f}(\xi) \in L_2$. Furthermore, for arbitrary real s , the space H^s can be defined as the set of all elements of \mathcal{S}' whose Fourier transform $\hat{f}(\xi)$ satisfies $(1 + |\xi|)^s \hat{f}(\xi) \in L_2$. H^s and H^{-s} are dual to each other.

For $f \in L_1$, $g \in L_2$, or $f, g \in L_2$, the convolution makes sense as a function, and it holds that $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$. This relation can be extended to distributions, to state

$$\mathcal{F}(S * T) = (\mathcal{F}S)(\mathcal{F}T).$$

This holds for $(S, T) \in \mathcal{S}' \times \mathcal{S}'$, $\mathcal{D}'_{L_2} \times \mathcal{D}'_{L_2}$ is the dual of \mathcal{D}_{L_2} , etc.

Fourier transforms are also often defined by

$$\hat{f}(\xi) = \int e^{-ix\xi} f(x) dx \quad \text{or} \quad = \int e^{-2\pi i x \xi} f(x) dx.$$

In the former case, the inversion formula takes the form

$$f(x) = (2\pi)^{-n} \int e^{ix\xi} \hat{f}(\xi) d\xi,$$

and the Parseval identity becomes

$$\int |f(x)|^2 dx = (2\pi)^{-n} \int |\hat{f}(\xi)|^2 d\xi.$$

References

- [1] N. Wiener, The Fourier integral and certain of its applications, Cambridge Univ. Press, 1933.
 - [2] S. Bochner, Vorlesungen über Fouriersche Integrale, Akademische-Verlag, 1932; English translation, Lectures on Fourier integrals, Ann. Math. Studies, Princeton Univ. Press, 1959.
 - [3] E. C. Titchmarsh, Introduction to the theory of Fourier integrals, Clarendon Press, 1937.
 - [4] R. E. A. C. Paley and N. Wiener, Fourier transforms in the complex domain, Amer. Math. Soc. Colloq. Publ., 1934.
 - [5] S. Bocher and K. Chandrasekharan, Fourier transforms, Ann. Math. Studies, Princeton Univ. Press, 1949.
 - [6] M. J. Lighthill, Introduction to Fourier analysis and generalized functions, Cambridge Univ. Press, 1958.
 - [7] I. M. Gel'fand and G. E. Shilov (Šilov), Generalized functions I, Academic Press, 1964. (Original in Russian, 1958.)
 - [8] S. Bochner, Harmonic analysis and the theory of probability, Univ. of California Press, 1955.
 - [9] R. R. Goldberg, Fourier transforms, Cambridge Tracts, Cambridge Univ. Press, 1961.
 - [10] T. Carleman, L'intégrale de Fourier et questions qui s'y rattachent, Almqvist and Wiksell, 1944.
 - [11] L. Schwartz, Théorie des distributions, Hermann, 1966.
- For formulas for Fourier transforms \rightarrow references to 220 Integral Transforms.

161 (IV.3) Free Groups

A. General Remarks

A group F is called a **free group** if it is the *free product (\rightarrow 190 Groups M) of *infinite cyclic groups G_1, \dots, G_n generated by a_1, \dots, a_n , respectively. Then n is called the **rank** of F . A

free product of \dagger semigroups is defined similarly to that of groups, and the free product of infinite cyclic semigroups $G_i = \{1, a_i, a_i^2, \dots\}$ ($i = 1, \dots, n$) is called a **free semigroup** generated by n elements a_i ($i = 1, \dots, n$).

If a group G is generated by subgroups H_i ($i = 1, \dots, n$) isomorphic to G_i , then G is a homomorphic image of the free product of the groups G_i . A subgroup $\neq \{e\}$ of the free product F of groups G_i is itself the free product of a free group and several subgroups, each of which is conjugate in F to a subgroup of some G_j (A. G. Kurosh, 1934). Notably, a subgroup $\neq \{e\}$ of a free group is itself a free group (O. Schreier, *Abh. Math. Sem. Univ. Hamburg*, 5 (1927)). A subgroup of index j of a free group of rank n is a free group of rank $1 + j(n-1)$ (Schreier).

Let F be the free group generated by n elements a_1, \dots, a_n , and let G be a group generated by n elements b_1, \dots, b_n . Then there is a homomorphism of F onto G . Let N be its kernel. If the class of a \dagger word $w(a_1, \dots, a_n)$ belongs to N , then we have $w(b_1, \dots, b_n) = 1$. We call $w(b_1, \dots, b_n) = 1$ a **relation** among the generators b_1, \dots, b_n . If N is the minimal normal subgroup of F containing the classes of words $w_1(a_1, \dots, a_n), \dots, w_m(a_1, \dots, a_n)$, then the relations $w_1(b_1, \dots, b_n) = 1, \dots, w_m(b_1, \dots, b_n) = 1$ are called **defining relations** (or **fundamental relations**). If generators a_1, \dots, a_n and words $w_1(a_1, \dots, a_n), \dots, w_m(a_1, \dots, a_n)$ are given, then there is a group generated by a_1, \dots, a_n with defining relations $w_1(a_1, \dots, a_n) = 1, \dots, w_m(a_1, \dots, a_n) = 1$. In fact, let F be the free group generated by a_1, \dots, a_n and N the minimal normal subgroup containing the classes of words $w_1(a_1, \dots, a_n), \dots, w_m(a_1, \dots, a_n)$. Then the factor group F/N is such a group. A free group is a group with an empty set of defining relations. In the preceding discussion, n and m are not necessarily finite. If both n and m are finite, then G is called **finitely presented**.

B. The Word Problem

If a finitely presented group G is given, then a general procedure has to be determined by which it can be decided, in a finite number of computational steps, whether a given word equals the identity element as an element of G . This is called the **word problem** (\rightarrow 190 Groups M). A solution to the word problem does not always exist (P. S. Novikov [8], 1955); in fact, there is a group with two generators and 32 defining relations for which the word problem cannot be solved (W. Boone [7]). However, it was shown by V. A. Tartakovskii that the problem can be solved for a large class of

groups. W. Magnus (1931) showed that it is solvable for any group with a single defining relation. The word problem is an example of decision problems (\rightarrow 97 Decision Problem). The word problem for groups is closely related to that for semigroups (A. M. Turing, 1937; E. L. Post, 1947; A. A. Markov, 1947). Similar problems for other algebraic systems can also be considered. The problem of determining a general procedure by which it can be decided, in a finite number of steps, whether two given words interpreted as elements of G can be transformed into each other by an (inner) automorphism of G is called the **transformation problem**.

Let F be a free group of rank n and $F = F_1 \supset \dots \supset F_r \supset F_{r+1} \supset \dots$ be the \dagger lower central series of F . Then F_r/F_{r+1} is a \dagger free Abelian group of rank $\mu_n(r) = (1/r) \sum_{d|r} \mu(r/d) n^d$, where μ is the \dagger Möbius function (E. Witt). The intersection of all subgroups of F of finite index is the identity element.

C. The Burnside Problem

The original problem of Burnside is: If every element of a group G is of finite order (but not necessarily of bounded order) and G is finitely generated, is G a finite group? E. S. Golod [6] (1964) showed that this problem for p -groups has a negative solution. The following is the more usual form of the **Burnside problem**: If a group G is finitely generated and the orders of elements of G divide a given integer r , is G finite? Let F be a free group of rank n , N be the normal subgroup of F generated by all the r th powers x^r of elements of F , and $B(r, n) = F/N$. Then the problem is the same as the question of whether $B(r, n)$ is finite. For $r = 2, 3, 4, 6$ the group is certainly finite (I. N. Sanov, M. Hall). The **restricted Burnside problem** is the question whether the orders of finite factor groups of $B(r, n)$ are bounded. It was solved affirmatively for r a prime (A. I. Kostrikin [5], 1959).

A group generated by two generators x, y and satisfying the relations $x^u = y^v = (xy)^w = 1$ (where u, v, w are integers) is infinite if $1/u + 1/v + 1/w - 1 \leq 0$, and is of order g if $0 < 1/u + 1/v + 1/w - 1 = 2/g$.

There is also a finitely presented group which is isomorphic to its proper factor group (B. H. Neumann).

References

- [1] A. G. Kurosh, The theory of groups I, II, Chelsea, 1960. (Original in Russian, 1953.)
- [2] M. Hall, The theory of groups, Macmillan, 1959.

- [3] H. S. M. Coxeter and W. O. J. Moser, *Generators and relations for discrete groups*, Erg. Math., Springer, 1957.
- [4] W. Magnus, A. Karrass, and D. Solitar, *Combinatorial group theory*, Interscience, 1966.
- [5] A. I. Kostrikin, *The Burnside problem*, Amer. Math. Soc. Transl., (2) 36 (1964), 63–100. (Original in Russian, 1959.)
- [6] E. S. Golod, *On nil-algebras and finitely approximable p -groups*, Amer. Math. Soc. Transl., (2) 48 (1965), 103–106. (Original in Russian, 1964.)
- [7] W. W. Boone, *The word problem*, Ann. of Math., (2) 70 (1959), 207–265.
- [8] P. S. Novikov, *On the algorithmic unsolvability of the word problem in group theory*, Amer. Math. Soc. Transl., (2) 9 (1958), 1–122. (Original in Russian, 1955.)
- Also → references to 190 Groups.

162 (XII.1) Functional Analysis

The origin of **functional analysis** can be traced to the 1887 work of V. Volterra. He stressed the important notion of operations or **operators**, that is, generalized functions of which the domains as well as the ranges are sets of functions. A typical example is the operator assigning to a function f its derivative f' . An operator is called a **functional** if its values are numbers, as in the case of the operator assigning to a function f the value $f'(a)$ or the value $\int_a^b f(t) dt$. In 1896, Volterra considered the operator mapping a continuous function f to a continuous solution φ of the integral equation $f(x) = \varphi(x) - \int_a^x K(x, y) \varphi(y) dy$, where $K(x, y)$ is a continuous function. Defining $I\varphi = \varphi$ and $(K\varphi)(x) = \int_a^x K(x, y) \varphi(y) dy$, he showed that φ is given by $\varphi = (I - K)^{-1}f = f + Kf + K^2f + \dots$, where $K^n f = K(K^{n-1}f)$. Following this lead, I. Fredholm studied in 1900 the integral equation $f(x) = \varphi(x) - \lambda \int_a^b K(x, y) \varphi(y) dy$ containing a parameter λ . He proved the so-called **†alternative theorem**: For a given λ_0 , the operator equation $(I - \lambda_0 K)\varphi = f$, $(K\varphi)(x) = \int_a^b K(x, y) \varphi(y) dy$, either admits a uniquely determined continuous solution φ for every continuous function f or else $(I - \lambda_0 K)\varphi_0 = 0$ admits a nontrivial continuous solution $\varphi_0 \neq 0$. D. Hilbert discussed (1904–1910) a **†continuous linear operator** K defined on the **†Hilbert space** L_2 with values in L_2 , and he called a complex number λ_0 a **†spectrum** of K if $(I - \lambda_0 K)$ does not have a continuous linear inverse. He proved that if K is **†Hermitian**, then K admits a **†spectral resolution** with real spectra only. One of his

outstanding contributions was the discovery of the **†continuous spectrum**. In 1918, F. Riesz proved that Fredholm's alternative theorem holds for a **†completely continuous linear operator** in the space of continuous functions, and later the result was extended to Banach spaces.

In 1932, three important books by S. Banach [1], J. von Neumann [2], and M. H. Stone [3] were published. These books treated **†closed linear operators** that are not necessarily continuous. The notion of **†Banach space** was introduced: a **†normed linear space** complete with respect to the distance $\text{dis}(x, y) = \|x - y\|$. By making use of the **†Baire-Hausdorff theorem** and the **†Hahn-Banach theorem**, Banach proved, for closed linear operators in Banach spaces, the fundamental principle consisting of the **†open mapping theorem**, the **†closed graph theorem**, the **†uniform boundedness theorem**, the **†resonance theorem**, and the **†closed range theorem**. These theorems were modified to be applicable in locally convex topological linear spaces by the Bourbaki group beginning in the late 1940s.

In 1929, von Neumann proved, as a mathematical foundation of quantum mechanics, that a closed linear operator T in a Hilbert space admits spectral resolution with real spectra if and only if T is a **†self-adjoint operator** (→ also Stone [3]). The condition that a closed linear operator is a **†function of a self-adjoint operator** was given by von Neumann, F. Riesz, and Y. Mimura (1934–1936). K. Friedrichs (1934) proved that a **†semibounded linear operator** admits a self-adjoint extension. T. Kato [4] (1950) proved that a Schrödinger-type Hermitian operator is **†essentially self-adjoint**.

Von Neumann's **†mean ergodic theorem** in the Hilbert space (1932) was extended to Banach spaces by K. Yosida, S. Kakutani, and Riesz in 1936. G. D. Birkhoff's **†pointwise ergodic theorem** (1931) was extended by N. Wiener (1939), Yosida (1940), E. Hopf (1954), N. Dunford (1955), R. V. Chacon and D. S. Ornstein (1960), and others in various ways. The **†Abelian ergodic theorems** were discussed by E. Hille and R. S. Phillips [5] and Yosida [6].

The notion of **†Banach algebra** was introduced by M. Nagumo in 1936. I. M. Gel'fand proved that a commutative Banach algebra with multiplicative unit e satisfying $\|e\| = 1$ (= the normed ring) admits a representation by an algebra of complex-valued continuous functions (1941).

The **†reflexivity** as well as the **†duality** of Banach spaces were studied by S. Kakutani (1939), V. L. Shmul'yan (1940), and W. T. Eberlein (1941).

The notion of \dagger vector lattice ($=\dagger$ Riesz space) was introduced in analysis by Riesz (1930). This was followed by the work of L. V. Kantorovich (1935) and H. Freudenthal (1936). Kakutani gave two standard types of \dagger Banach lattice: the \dagger abstract (M) space (1940) and the \dagger abstract (L) space (1941). M. Krein and S. Krein (1940) and Yosida and M. Fukamiya (1940) discussed the (M)-type vector lattices, and Yosida (1941), the (L)-type vector lattices. H. Nakano (1940–1941) and T. Ogasawara and F. Maeda (1942) studied the spectral resolution of the Banach lattice.

The connection between \dagger Brownian motion and \dagger potential theory was clarified by Kakutani (1942), and extended by J. L. Doob (1956) and G. A. Hunt (1957–1958).

The \dagger one-parameter semigroup of continuous linear operators in Banach spaces was studied by Hille and Yosida, and they gave in 1948 a characterization of the \dagger infinitesimal generator of such semigroups. Its dual was given by Phillips (1955). The one-parameter semigroup of nonlinear \dagger contractive operators in Hilbert spaces was studied by Y. Kōmura (1967), who obtained a nonlinear version of the Hille-Yosida theorem. This has been extended considerably in Banach spaces by many scholars, e.g., Kato (1967), M. G. Crandall and A. Pazy (1969), Crandall and T. Liggett (1971), I. Miyadera and S. Oharu (1970), H. Brezis (1973), P. Bénéilan (1973), and others.

In 1936, S. Sobolev gave a generalization of the notion of functions and their derivatives through integration by parts. This generalization has been extended by L. Schwartz (1945–) [9] to the notion of \dagger distributions, which are continuous linear functionals defined on the function spaces $\mathcal{D}(\mathbf{R}^n)$ and $\mathcal{S}(\mathbf{R}^n)$, and this extension gives, e.g., a reasonable interpretation of Dirac's δ -function. Since 1959, Gel'fand [10] has been publishing, with his collaborators, books on the distribution theory pertaining to function spaces other than $\mathcal{D}(\mathbf{R}^n)$ and $\mathcal{S}(\mathbf{R}^n)$. M. Sato introduced (1959–1960) [11] the theory of \dagger hyperfunctions, as a generalization of distribution theory. In the case of one independent variable, a hyperfunction f may be defined as a generalized boundary value on the real axis R of a holomorphic function F defined in $\mathbf{C}^1 - \mathbf{R}^1$. Hyperfunction theory has been refined to \dagger microlocal analysis and studied extensively by Sato, A. Martineau, H. Komatsu, P. Schapira, T. Kawai, M. Kashiwara, M. Morimoto, A. Kaneko, and others (\rightarrow [12]).

References

- [1] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
- [2] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, 1932.
- [3] M. H. Stone, *Linear transformations in Hilbert spaces and their applications to analysis*, Amer. Math. Soc., 1932.
- [4] T. Kato, *Perturbation theory for linear operators*, Springer, second edition, 1976.
- [5] E. Hille and R. S. Phillips, *Functional analysis and semigroups*, Amer. Math. Soc., 1957.
- [6] K. Yosida, *Functional analysis*, Springer, sixth edition, 1980.
- [7] N. Dunford and J. Schwartz, *Linear operators*, Interscience, I, 1958; II, 1963; III, 1971.
- [8] H. Brezis, *Opérateurs Maximaux Monotones et Semigroupes de Contractions dans les Espaces de Hilbert*, American Elsevier, 1973.
- [9] L. Schwartz, *Théorie des Distributions*, Hermann, 1966.
- [10] I. M. Gel'fand, *Generalized functions I–III* (with G. E. Shilov), IV (with N. Ya. Vilenkin), V (with M. I. Graev and N. Ya. Vilenkin), Academic Press, 1964–1966.
- [11] M. Sato, *Theory of Hyperfunctions I, II*, J. Fac. Sci. Univ. Tokyo, sec. I, 8 (1959), 139–193; 8 (1960), 387–437.
- [12] M. Sato, T. Kawai, and M. Kashiwara, *Microfunctions and pseudodifferential equations*, Lecture notes in math. 287, Springer, 1973.

163 (XIII.16) Functional-Differential Equations

A. General Remarks

In many models, it is assumed that the future behavior of a system under consideration is governed only by its present state and not by past states. However, for various systems arising in practical problems we cannot ignore the effect of the past on the future. Such a phenomenon is often observed in population problems, epidemiology, chemical reactions, system engineering, and so on.

The description of such phenomena may involve **difference-differential equations**

$$\dot{x}(t) = f(t, x(t), x(t-h_1), \dots, x(t-h_m)), \quad (1)$$

or **integro-differential equations**

$$\dot{x}(t) = g(t, x(t)) + \int_0^t f(t, s, x(t), x(s)) ds. \quad (2)$$

An enormous variety of equations is discussed in the literature, but most of them can be

expressed in the form

$$\dot{x}(t) = f(t, x(s)), \quad (3)$$

which is usually called the **functional differential equation**; here s varies in an interval I_t and f depends on the function x of I_t . If t is the right or left endpoint of the interval I_t , then (3) is said to be of **retarded** or **advanced type**, respectively. Most of the results obtained have been from work on equations of retarded type. In this case, the maximal length of the interval I_t is called the **retardation (delay, lag, deviation, etc.)**, and (3) is often called the **retarded differential equation, delay differential equation, or differential equation with lag (retardation, deviating argument)**.

B. Historical Remarks

Functional differential equations of a sort were studied by Johann Bernoulli in the 18th century in connection with the string problem. Since then, a good deal of work has been done in this field by many people; some of this work was done before the beginning of this century. Among the investigators, Volterra [1, 2] is noteworthy for his systematic study on rather general equations related to problems of predator-prey populations and viscoelasticity, though his results were ignored by his contemporaries. In the early 1940s, Minorsky [3], in his famous study of ship stabilization, pointed out the importance of delay effects in control theory, with many of the modern issues first appearing in his work [4]. Mishkis [5] studied linear systems extensively, and Driver [6] gave a unified representation for functional differential equations. Important achievements were given by Krasovskii [7] and Bellman and Cooke [8]. They laid the foundation for the qualitative theory of functional differential equations. Inspired by these works, many books (such as [9–14]) were published. Owing to these books together with many articles on this field, the theory of functional differential equations has become an important branch in the theory of differential equations.

Equations of advanced type are treated in, e.g., [5, 8], but qualitative theory for them has hardly been established. The equation

$$\dot{x}(t) = ax(t) + bx(\lambda t) \quad \text{on} \quad 0 \leq t < \infty \quad (4)$$

is of advanced type if $\lambda > 1$, and its analytic solution has been studied in detail. Equation (3), whose right-hand side involves differential operators, e.g., $\dot{x}(t) = f(t, x(t), x(t-h), \dot{x}(t-h))$, is said to be of **neutral type** [5, 8, 14], and is generally considered to be of neither retarded nor advanced type because of its distinctive

features. However, Hale [14] originated the study of a class of equations of neutral type for which the general qualitative theory can be developed in the same manner as for equations of retarded type.

For the case of infinite retardation, proper choice of the space of functions x contained in the right-hand side of (3) is of great significance. A general treatment of the phase space for equations with infinite retardation in an axiomatic setting is given in [15].

C. Phase Space

For simplicity, let I_t always be the interval $[t-h, t]$ with a finite retardation $h > 0$. A **solution** of (3) starting at $t = \tau$ is a function defined on $[\tau-h, a)$ for $a > \tau$ such that the solution coincides with a preassigned function (the **initial function**) on $[\tau-h, \tau]$ and is continuously differentiable and satisfies (3) on $[\tau, a)$. Since a solution starting at $t = \tau$ must be continuous for $t \geq \tau$, it is quite natural, as is done in much of the literature in order to develop a qualitative theory, to choose the space $C([t-h, t], \mathbf{R}^n)$ of continuous \mathbf{R}^n -valued functions on $[t-h, t]$ as a space of functions x involved in the right-hand side of (3). Introducing a symbol x_t which represents an element of $C = C([-h, 0], \mathbf{R}^n)$ defined by $x_t(s) = x(t+s)$ ($s \in [-h, 0]$), we can rewrite (3) in a more convenient form:

$$\dot{x}(t) = f(t, x_t), \quad (E) \quad (5)$$

where the function $f(t, \varphi)$ is defined on $\mathbf{R} \times C$ (or on its subspace). Here C is considered to be a Banach space with the norm $\|\varphi\| = \max_{s \in [-h, 0]} |\varphi(s)|$, $|\cdot|$ being a norm in \mathbf{R}^n . The space C is said to be the **phase space** of (E), and the initial condition can be written as

$$x_\tau = \xi, \quad (\tau, \xi) \in \mathbf{R} \times C. \quad (5)$$

D. Initial Value Problem

Let $f(t, \varphi)$ be a continuous functions defined on an open domain $D \subset \mathbf{R} \times C$. The **initial value problem** is to find a solution of (E)–(5) for a given $(\tau, \xi) \in D$. Many fundamental results hold as for ordinary differential equations: (a) There always exists a solution of (E)–(5) under the continuity of f on D . Here, the solution means the one to the right. No general existence theorem is given for the solution to the left. (b) The solution of (E)–(5) is unique, if f satisfies the **Lipschitz condition**: $|f(t, \varphi) - f(t, \psi)| \leq L \|\varphi - \psi\|$ in a neighborhood of each point of D , where L is a constant depending on the neighborhood. (c) If f varies in the space $C(D, \mathbf{R}^n)$

equipped with the compact open topology and if the solution of (E)–(5) is unique for every $(\tau, \xi) \in D$, then the solution is continuous as a function of (τ, ξ, f) . (d) If f is completely continuous on D , that is, f is a continuous mapping of bounded subsets of D into bounded sets in \mathbf{R}^n , then every solution can be continuable to the right as long as it remains bounded or stays away from the boundary of D . This assertion is no longer true in the absence of complete continuity.

To prove the existence theorem, Schauder's fixed point theorem is utilized; the Picard successive method is also effective under the Lipschitz condition. Consider the case when $f(t, \varphi)$ can be written as $f(t, \varphi) = g(t, \varphi(0), \varphi)$, where $g(t, x, \varphi)$ is defined for $(t, x, \varphi) \in \mathbf{R} \times \mathbf{R}^n \times C$, and $g(t, x, \varphi) = g(t, x, \psi)$ if $\varphi(s) = \psi(s)$ on $[-h, -\delta]$, $\delta > 0$. Equation (1) is one such case, where $h = \max h_k$, $\delta = \min h_x$. Then, $g(t, x, x_t)$ is a function of (t, x) alone on $[\tau, \tau + \delta]$ under the given condition (5), that is, (E) is reduced to an ordinary differential equation. This makes it possible to find a solution of (E)–(5) by matching successively the solutions of ordinary differential equations on $[\tau, \tau + \delta]$, $[\tau + \delta, \tau + 2\delta]$, This is the **step-by-step method**, which is effective even for equations of neutral type with the same property.

Under uniqueness, the solution $x(t)$ of (E) induces a mapping $T(t, \tau): G(t, \tau) \rightarrow C$, $t \geq \tau$, which maps $x_t \in G(t, \tau)$ to $x_t \in C$, where $G(t, \tau) \subset C$ is the set of ξ for which the solution of (E)–(5) is continuable up to t . Then

$$T(t, s)T(s, \tau) = T(t, \tau) \quad \text{for} \quad t \geq s \geq \tau, \tag{6}$$

and $T(t, \tau)$ is strongly continuous in t . If (E) is **autonomous**, that is, $f(t, \varphi)$ is independent of t , then there exists a mapping $T(t)$, $t \geq 0$, satisfying $T(t - \tau) = T(t, \tau)$, and $\{T(t)\}_{t \geq 0}$ becomes a one-parameter semigroup.

E. Linear System

When $f(t, \varphi)$ is continuous on $I \times C$ for an interval I and is linear in φ , equation (E) is said to be a **linear system** (denoted by (L)). In this case, $f(t, \varphi)$ satisfies the Lipschitz condition with $L = L(t)$ on $I \times C$ for a continuous function $L(t)$ of I , which also implies that f is completely continuous. Thus the solution of (L)–(5) uniquely exists over $[\tau, \infty) \cap I$ for any $(\tau, \xi) \in I \times C$, and the mapping (the **fundamental** or **solution operator**) $T(t, \tau)$ is a \dagger bounded linear operator on C for any $t, \tau \in I$, $t \geq \tau$, and \dagger compact if $t \geq \tau + h$. $T(t, \tau)$ corresponds to the fundamental matrix for ordinary linear differential equations, but it is not invertible in general.

A continuous function $f(t, \varphi)$ linear in φ

may be expressed in the form

$$f(t, \varphi) = \int_{-h}^0 [d_s \eta(t, s)] \varphi(s), \tag{7}$$

where $\eta(t, s)$ is an $n \times n$ matrix defined on $I \times [-h, 0]$, which is measurable in (t, s) and of \dagger bounded variation in s with a \dagger total variation less than $L(t)$, and the integration is that of Stieltjes with respect to s . From this fact it follows that the range of the initial functions can be extended to the space of \dagger piecewise-continuous functions, and the solution $x(t)$ of the nonhomogeneous linear system $\dot{x}(t) = f(t, x_t) + p(t)$ through $(\tau, \xi) \in I \times C$ can be represented by the **constant variational formula**

$$x(t) = [T(t, \tau)\xi](0) + \int_{\tau}^t [T(t, s)\Gamma](0)p(s)ds, \quad t \geq \tau,$$

where $\Gamma(s) = E$ (the unit matrix) for $s = 0$, and $\Gamma(s) = O$ (the zero matrix) for $s < 0$.

F. Autonomous Linear System

Let (L) be an autonomous linear system

$$\dot{x}(t) = f(x_t). \tag{AL}$$

In this case, (7) is not different from the Riesz representation theorem $f(\varphi) = \int_{-h}^0 [d\eta(s)] \varphi(s)$. For (AL) the fundamental operator $T(t)$ plays an extremely important role. As was seen, $\{T(t)\}_{t \geq 0}$ is a one-parameter semigroup of bounded linear operators $T(t)$ which is strongly continuous in $t \geq 0$ and compact for $t \geq h$. Thus the asymptotic behavior of the solutions of (AL) are determined by the distribution of the spectra $\sigma(A)$ of the infinitesimal generator A of $T(t)$, which is given by $A\varphi = \dot{\varphi}$ with the domain $D(A) = \{\varphi \in C \mid \dot{\varphi} \in C \text{ and } \varphi(0) = f(\varphi)\}$. The properties of $T(t)$ assert that: (a) $\sigma(A)$ consists of point spectra alone. (b) The number of $\Lambda_\alpha = \{\lambda \in \sigma(A) \mid \operatorname{Re} \lambda \geq \alpha\}$ is at most finite for any $\alpha \in \mathbf{R}$. (c) For every $\lambda \in \sigma(A)$ the dimension of the generalized eigenspace of λ is finite. (d) The spectra of A coincide with the roots (**characteristic roots**) of the **characteristic equation** of (AL),

$$\det \left[\lambda E - \int_{-h}^0 e^{\lambda s} d\eta(s) \right] = 0,$$

together with their multiplicities. (e) $e^{\lambda t} \in \sigma(T(t))$, $t \geq 0$, if and only if $\lambda \in \sigma(A)$.

Let P_α , $\alpha \in \mathbf{R}$, be the linear space spanned by the generalized eigenfunctions corresponding to a $\lambda \in \Lambda_\alpha$. Then, (a) P_α is invariant under $T(t)$, that is, $T(t)P_\alpha \subset P_\alpha$ for $t \geq 0$; (b) the restriction of $T(t)$ to P_α is invertible and hence extendable over $t \in \mathbf{R}$; (c) if $\xi \in P_\alpha$, then $[T(t)\xi](0) = \sum_{\lambda \in \Lambda_\alpha} p_{\lambda}(t, \xi)e^{\lambda t}$, where $p_{\lambda}(t, \xi)$ are polynomials

in t and linear in ξ ; (d) there is a direct-sum decomposition $C = P_\alpha + Q_\alpha$ such that Q_α is invariant under $T(t)$, the projection along $Q_\alpha \pi_\alpha: C \rightarrow P_\alpha$ is bounded, and there exist positive constants K, ε for which $\|T(t)\xi\| \leq K e^{(\alpha - \varepsilon)t} \|\xi\|$, $t \geq 0$ if $\xi \in Q_\alpha$. Hence for any $\xi \in C$ the solution $x(t)$ of (AL) through ξ at $t=0$ satisfies

$$\left| x(t) - \sum_{\lambda \in \Lambda_\alpha} p_\lambda(t, \pi_\alpha \xi) e^{\lambda t} \right| \leq K(1 + \|\pi_\alpha\|) e^{(\alpha - \varepsilon)t} \|\xi\|, \quad t \geq 0, \quad (8)$$

where $\|\pi_\alpha\|$ denotes the operator norm of π_α . The set $Q = \bigcap_{\alpha \in R} Q_\alpha$ may contain an element ξ other than the zero element, but $T(t)\xi$ must be identically zero for $t \geq h$.

For a linear system (L) with $f(t, \varphi)$ ω -periodic in t , similar conclusions result, with $T(\omega, 0)$ in place of A ; and relation (8) holds, where Λ_α is defined similarly by replacing $\sigma(A)$ by $\{(1/\omega) \log \mu \mid \mu \in \sigma(T(\omega, 0))\}$ and $p_\lambda(t, \xi)$ are polynomials with ω -periodic coefficients. This corresponds to Floquet's theorem for ordinary periodic linear systems. $\mu \in \sigma(T(\omega, 0))$ and $(1/\omega) \log \mu$ are said to be the **characteristic multiplier** and the **characteristic exponent**, respectively.

G. Stability Problem

The concept of stability can be defined and studied in the same spirit as for ordinary differential equations, and the Lyapunov second method also turns out to be effective. For instance, the zero solution of (E) with f defined on $D_H = [0, \infty) \times \{\varphi \in C \mid \|\varphi\| < H\}$ for $H > 0$ is said to be **uniformly asymptotically stable** if there are a constant $\alpha > 0$ and positive functions $\delta(\varepsilon)$ and $\sigma(\varepsilon)$ of $\varepsilon > 0$ such that any solution $x(t)$ of (E) satisfies

$$|x(t)| < \varepsilon \text{ as long as } x(t) \text{ exists,} \quad (9)$$

whenever $\|x_\tau\| < \delta(\varepsilon)$ and $t \geq \tau$ or $\|x_\tau\| \leq \alpha$ and $t \geq \tau + \sigma(\varepsilon)$ for $\tau \geq 0$. In the above definition, if f is completely continuous on D_H and if $0 < \varepsilon < H$, then the phrase "as long as $x(t)$ exists" is redundant, and (9) is equivalent to the claim that $x(t)$ exists for all $t \geq \tau$ and $|x(t)| < \varepsilon$. Under the Lipschitz condition and the complete continuity on f in (E), the zero solution of (E) is uniformly asymptotically stable if and only if there exists a continuous \mathbf{R} -valued function (**Lyapunov function**) $V(t, \varphi)$ defined on D_{H_1} for $0 < H_1 \leq H$ such that

$$(i) \quad a(\|\varphi(0)\|) \leq V(t, \varphi) \leq b(\|\varphi\|),$$

$$(ii) \quad D^+ V(t, x_t) \leq -cV(t, x_t)$$

as long as $(t, x_t) \in D_{H_1}$ for a solution $x(t)$ of (E), where $a(r), b(r)$ are continuous functions with $a(r) > 0$ for $r > 0$, $b(0) = 0$, $c > 0$ is a constant,

and D^+ denotes the upper-right Dini derivative. Furthermore, the Lyapunov function can be endowed with a Lipschitz condition $|V(t, \varphi) - V(t, \psi)| \leq L_1 \|\varphi - \psi\|$ for some constant L_1 if f in (E) is uniformly Lipschitzian, that is, the coefficient L is constant over the domain. The Lipschitz condition together with (ii) makes it possible for a solution $y(t)$ of the perturbed equation

$$\dot{y}(t) = f(t, y_t) + P(t, y_t) \quad (10)$$

to satisfy

$$D^+ V(t, y_t) \leq -cV(t, y_t) + L_1 |P(t, y_t)|.$$

By using this fact, the stability of (10) can be discussed.

However, for functional differential equations it is quite a difficult problem to construct a suitable Lyapunov function for a given (E).

Many of the attempts to improve the sufficiency condition that have been made up to now are such that a stability property can be verified by means of a simple Lyapunov function. The main effort has been devoted to replacing condition (ii) by another type of condition. One of them is

$$(ii^*) \quad D^+ V(t, x_t) \leq -c(|x(t)|)$$

under the uniform boundedness of $f(t, \varphi)$, where $c(r)$ is a continuous function with $c(r) > 0$ for $r > 0$. Another one is for the case when $V(t, \varphi) = W(t, \varphi(0))$, where $W(t, x)$ is defined for $(t, x) \in \mathbf{R} \times \mathbf{R}^n$. In this case, (ii) can be replaced by

$$(ii^{**}) \quad D^+ V(t, x_t) \leq -cV(t, x_t)$$

whenever $V(t + s, x_{t+s}) \leq F(V(t, x_t))$ for $s \in [-h, 0]$, where $c > 0$ is a constant and $F(r)$ is a continuous function satisfying $F(r) > r$ for $r > 0$. The condition (ii**) was given by B. S. Razumikhin and provides an easier way to construct a Lyapunov function.

For a linear system, uniform asymptotic stability implies $\|T(t, \tau)\| \leq 1/\delta(1)$ for $t \geq \tau$ and $\|T(t, \tau)\| \leq 1/2$ for $t \geq \tau + \sigma(\alpha/2) + h$. These facts together with (6) show that the zero solution is **exponentially stable**, that is, $|x(t)| \leq K e^{-\nu(t-\tau)} \|x_\tau\|$ for $t \geq \tau$, where $\nu = (\log 2)/(\sigma(\alpha/2) + h)$, $K = 2/\delta(1)$. From (8) it follows that the zero solution of (AL) is uniformly asymptotically stable if and only if

$$\{\lambda \in \sigma(A) \mid \operatorname{Re} \lambda \geq 0\} = \emptyset. \quad (11)$$

H. Equations of Neutral Type

To deal with an equation of neutral type, such as

$$\dot{x}(t) = f(t, x_t, \dot{x}_t), \quad (12)$$

a phase space such as $C^1 = C^1([-h, 0], \mathbf{R}^n)$ of continuously differentiable functions should be chosen. The solution of (E) will become smooth as time elapses if f in (E) is sufficiently smooth. However, this property cannot be expected for (12), and the solution of (12)–(5) need not belong to C^1 even if $\xi \in C^1$ without a specific condition on ξ such as $\dot{\xi}(0) = f(\tau, \xi, \dot{\xi})$. Such a restriction on initial functions is an obstruction to the development of the general theory. Equations of the form

$$\frac{d}{dt}D(t, x_t) = f(t, x_t) \quad (\text{N})$$

cover a fairly general class of equations of neutral type, where $D(t, \varphi)$ and $f(t, \varphi)$ are defined on an open set $D \subset \mathbf{R} \times C$ and a **solution** of (N)–(5) means a continuous function $x(t)$ defined on $[\tau - h, a]$, $a > \tau$, which satisfies (5) and

$$D(t, x_t) = D(\tau, \xi) + \int_{\tau}^t f(s, x_s) ds \quad \text{on } [\tau, a].$$

(E) corresponds to the case where $D(t, \varphi) = \varphi(0)$, while $\dot{x}(t) = G(t, \dot{x}_t) + g(t, x_t)$ is reduced to the form (N) if $G(t, \varphi)$ is linear in φ and continuously differentiable with respect to t by setting $D(t, \varphi) = \varphi(0) - G(t, \varphi)$ and $f(t, \varphi) = g(t, \varphi) - (\partial/\partial t)G(t, \varphi)$.

A continuous function $D(t, \varphi)$ linear in φ is said to be **atomic at 0** if in the representation

$$D(t, \varphi) = \int_{-h}^0 [d_s \mu(t, s)] \varphi(s)$$

(see equation (7)) $P(t) = \mu(t, 0) - \mu(t, -0)$ exists and is nonsingular, which is equivalent to

$$D(t, \varphi) = P(t)\varphi(0) + \int_{-h}^0 [d_s \mu_0(t, s)] \varphi(s) \quad (13)$$

with a nonsingular matrix $P(t)$ and a matrix $\mu_0(t, s)$ whose total variation with respect to s on $[-\sigma, 0]$ tends to 0 as $\sigma \rightarrow 0$ locally uniformly in t . A nonlinear function $D(t, \varphi)$ is said to be **atomic at 0** if $D(t, \varphi)$ has a continuous † Fréchet derivative $D_{\varphi}(t, \varphi)$ with respect to φ and $D_{\varphi}(t, \varphi)$ is atomic at 0. In equation (N), $D(t, \varphi)$ is always assumed to be atomic at 0, and many of the fundamental theorems for (E) are also valid for (N). Let $T(t)$ be the operator solution of the autonomous linear equation

$$\frac{d}{dt}D(x_t) = f(x_t). \quad (14)$$

Then $\{T(t)\}_{t \geq 0}$ is a strongly continuous semi-group of bounded linear operators, and the corresponding † infinitesimal generator A has the domain $D(A) = \{\varphi \in C \mid \dot{\varphi} \in C \text{ and } D(\dot{\varphi}) = f(\varphi)\}$. The properties of the spectra $\sigma(A)$ are the same as for (AL), except that the character-

istic equation is now given by

$$\det \left[\lambda \left(P + \int_{-h}^0 e^{\lambda s} d\mu_0(s) \right) - \int_{-h}^0 e^{\lambda s} d\eta(s) \right] = 0$$

and the property (b) is true if $\alpha > a_D$, where $D(\varphi)$, $f(\varphi)$ are assumed to be of the forms (13) and (7), respectively, without the argument t , and where a_D is given by

$$a_D = \max \left\{ \operatorname{Re} \lambda \mid \det \left[P + \int_{-h}^0 e^{\lambda s} d\mu_0(s) \right] = 0 \right\}.$$

Similarly, the decomposition $C = P_{\alpha} + Q_{\alpha}$ is possible if $\alpha > a_D$. Hence, to obtain the uniform asymptotic stability of the zero solution of (14), the condition $a_D < 0$ should be assumed in addition to (11). The linear function $D(\varphi)$ is said to be **stable** if $a_D < 0$. If $D(\varphi)$ is stable, then the Lyapunov method is also applicable, as a sufficient condition, to the stability problem of (N) with linear $D(\varphi)$ instead of general $D(t, \varphi)$, but in this case condition (i) for $V(t, \varphi)$ should be replaced by

$$a(|D(\varphi)|) \leq V(t, \varphi) \leq b(\|\varphi\|).$$

I. Infinite Retardation

There are some basic differences between cases of finite retardation and of infinite retardation. For example, in the definition of stability, the inequality $|x(t)| < \varepsilon$ in (9) can be replaced by $\|x_t\| < \varepsilon$ with no difference in the case of finite retardation, but this replacement yields a different concept deeply connected with the choice of the phase space in case of infinite retardation.

There are several ways to define a phase space for a functional differential equation with infinite retardation. One of those phase spaces generally used is a linear space X of functions: $(-\infty, 0] \rightarrow \mathbf{R}^n$ with a seminorm $\|\cdot\|_X$ such that if a function $x: (-\infty, a) \rightarrow \mathbf{R}^n$ satisfies $x_t \in X$ and it is continuous on $[\tau, a]$, then

- (i) $x_t \in X$ for all $t \in [\tau, a]$,
- (ii) x_t is continuous as a function $[\tau, a) \rightarrow X$,
- (iii) $m|x(t)| \leq \|x_t\|_X \leq K \max_{s \in [\tau, t]} |x(s)| + M(t - \tau)\|x_t\|_X$,

where m and K are positive constants and $M(t)$ is a continuous function.

If X satisfies the foregoing conditions and if $f(t, \varphi)$ is continuous on an open domain $D \subset \mathbf{R} \times X$, then the local properties (a)–(c) in Section D hold, and so does (d) when $D = \mathbf{R} \times X$. On the other hand, if X has a **fading memory**, namely,

- (iv) $M(t) \leq M e^{-\mu t}$ in (iii) for positive constants M, μ ,

then the procedure in Section F is applicable by restricting $\sigma(A)$ to $\sigma_\mu(A) = \{\lambda \in \sigma(A) \mid \operatorname{Re} \lambda > -\mu\}$ and, hence the zero solution of (AL) with infinite retardation is uniformly asymptotically stable under (11).

For $0 \leq h \leq \infty$ and $0 \leq \gamma < \infty$, both the space $C_{h,\gamma}$ of continuous functions $\varphi: (-h, 0] \rightarrow \mathbf{R}^n$ with a finite limit $\lim_{s \rightarrow -h} e^{\gamma s} \varphi(s)$ and the space $M_{h,\gamma}$ of measurable functions $\varphi: (-h, 0] \rightarrow \mathbf{R}^n$ with $\int_{-h}^0 e^{\gamma s} |\varphi(s)| ds < \infty$ satisfy the foregoing conditions, where the norms are given by $\|\varphi\|_{C_{h,\gamma}} = \sup_{(-h, 0]} e^{\gamma s} |\varphi(s)|$ and $\|\varphi\|_{M_{h,\gamma}} = |\varphi(0)| + \int_{-h}^0 e^{\gamma s} |\varphi(s)| ds$. These spaces have a fading memory if $h < \infty$ or $\gamma > 0$.

In much of the literature in which the equation of infinite retardation has a form like (2) or (4), the space CB of bounded continuous functions or the space C_0 of continuous functions with compact supports are sufficient as a phase space under the norm $\|\varphi\| = \sup_{s \leq 0} |\varphi(s)|$. However, it is to be noted that C_0 is not complete when CB does not satisfy condition (ii). When CB is chosen as the phase space, in order to show the existence theorem, $f(t, x_t)$ should be continuous for any bounded continuous function x in addition to the continuity in (t, φ) . This condition is satisfied if $I_t = [g(t), t]$ in (3) for a continuous function $g(t) \leq t$. Equation (2) or (4) with $0 \leq \lambda < 1$ give rise to such a case, but $g(t)$ must be equal to 0 in (2) and (4) with $\lambda = 0$. If $g(t) \rightarrow \infty$ as $t \rightarrow \infty$, the Razumikhin condition (ii**) $D^+ V(t, x_t) \leq -cV(t, x_t)$ whenever $V(s, x_s) \leq F(V(t, x_t))$ ($s \in [g(t), t]$) for a Lyapunov function is effective, but one can conclude only that $x(t) \rightarrow 0$ as $t \rightarrow \infty$ without uniformity.

References

- [1] V. Volterra, Sur la théorie mathématique des phénomènes héréditaires, *J. Math. Pures Appl.*, 7 (1928), 249–298.
- [2] V. Volterra, *Théorie mathématique de la Lutte pour la Vie*, Gauthier-Villars, 1931.
- [3] N. Minorsky, Self-excited oscillations in dynamical systems possessing retarded actions, *J. Appl. Mech.*, 9 (1942), 65–71.
- [4] A. M. Zverkin, G. A. Kamenskii, S. B. Norkin, and L. E. El'sgol'ts, Differential equations with retarded arguments (in Russian), *Uspekhi Mat. Nauk.*, 17 (1962), 77–164.
- [5] A. D. Mishkis, *Lineare Differentialgleichungen mit nachteilendem Argument*, Deutscher Verlag Wiss., 1955. (Original in Russian, 1957.)
- [6] R. D. Driver, Existence and stability of solution of a delay-differential system, *Arch. Rational Mech. Anal.*, 10 (1962), 401–426.
- [7] N. N. Krasovskii, *Stability of motion*, Stanford Univ. Press, 1963. (Original in Russian, 1959.)
- [8] R. Bellman and K. L. Cooke, *Differential-difference equations*, Academic Press, 1963.
- [9] L. E. El'sgol'ts and S. B. Norkin, *Introduction to the theory and application of differential equations with deviating argument*, Academic Press, 1971. (Original in Russian, 1963.)
- [10] A. Halanay, *Differential equations: Stability, oscillations, time lags*, Academic Press, 1965.
- [11] M. N. Oğuztöreli, *Time-lag control systems*, Academic Press, 1966.
- [12] T. Yoshizawa, *Stability theory by Liapunov's second method*, Math. Soc. Japan, 1966.
- [13] V. Lakshmikantham and S. Leela, *Differential and integral inequalities II*, Academic Press, 1969.
- [14] J. K. Hale, *Theory of functional differential equations*, Springer-Verlag, 1977.
- [15] J. K. Hale and J. Kato, Phase space for retarded equations with infinite delay, *Funkcial. Ekvac.*, 21 (1978), 11–41.

164 (XII.18) Function Algebras

A. Definition [1–3]

Let $C(X)$ be the \dagger Banach algebra of all continuous complex-valued functions on a compact Hausdorff space X with pointwise operations and \dagger uniform norm. A **function algebra** (or **uniform algebra**) on X is a closed subalgebra A of $C(X)$ containing the constant functions and separating the points of X , i.e., for any $x, y \in X$, $x \neq y$, there exists $f \in A$ with $f(x) \neq f(y)$. If $\varphi_x, x \in X$, denotes the evaluation mapping $f \rightarrow f(x)$ of A , the correspondence $x \rightarrow \varphi_x$ is a homeomorphism of X into the \dagger maximal ideal space $\mathfrak{M}(A)$ of A . Since $\hat{f}(\varphi_x) = f(x)$ for any $x \in X$ and $f \in A$, the \dagger Gel'fand representation of A is an isometric isomorphism. By identifying x with φ_x , we regard X as a closed subset of $\mathfrak{M}(A)$ and the Gel'fand transform $\hat{f}, f \in A$, as a continuous extension of f to $\mathfrak{M}(A)$.

B. Examples [1, 3, 4]

For a compact plane set K let $P(K)$ ($R(K)$) be the subalgebra of all functions in $C(K)$ that can be approximated uniformly on K by polynomials (rational functions with poles off K). $A(K)$ denotes the subalgebra of all functions in $C(K)$ that are analytic in the interior of K . These are function algebras on K and $P(K) \subseteq R(K) \subseteq A(K)$. When K is the unit circle $\mathbf{T} =$

$\{|z|=1\}$, $P(\mathbf{T})$ is called the **disk algebra**, the most typical concrete example.

The theory of function algebras emerged from attempts to solve, by means of functional analysis, certain problems in complex analysis, especially problems of uniform approximation, e.g., when does $A(K)=P(K)$ or $A(K)=R(K)$ hold?

Another important example is given by the algebra $H_\infty(U)$ of all bounded analytic functions on a bounded plane domain U with the supremum norm $\|f\|_\infty = \sup\{|f(z)| \mid z \in U\}$ [5]. Since the Gelfand representation is an isometric isomorphism of $H_\infty(U)$ into the algebra $C(\mathfrak{M}(H_\infty(U)))$, $H_\infty(U)$ is viewed as a function algebra on the space $\mathfrak{M}(H_\infty(U))$. Further examples result if we take K or U in a Riemann surface or in the complex n -space.

We list a few abstract function algebras that reflect certain relevant properties possessed by concrete examples. A is called a **Dirichlet** (resp. **logmodular**) **algebra** if the set $\text{Re } A = \{\text{Re } f \mid f \in A\}$ (resp. $\log |A^{-1}| = \{\log |f| \mid f, f^{-1} \in A\}$) is dense in $C_{\mathbf{R}}(X)$, the space of all continuous real-valued functions on X . It is called **hypo-Dirichlet** if the closure of $\text{Re } A$ has finite codimension in $C_{\mathbf{R}}(X)$, and the linear span of $\log |A^{-1}|$ is dense in $C_{\mathbf{R}}(X)$.

In the following, A denotes a function algebra on X unless otherwise specified.

C. Boundary and Representing Measure [1–4]

A subset E of X is a **boundary** for A if for any $f \in A$ there exists $x \in E$ such that $|f(x)| = \|f\|$.

A **closed boundary** is a boundary closed in X . G. E. Shilov proved that there is a smallest closed boundary, ∂A , which is called the **Shilov boundary** for A . A positive Borel measure μ on X is a **representing measure** for $\varphi \in \mathfrak{M}(A)$ if $f(\varphi) = \int f(x) d\mu(x)$ for all $f \in A$. Each $\varphi \in \mathfrak{M}(A)$ has a representing measure supported by ∂A . The **Choquet boundary**, $c(A)$, consists of all $x \in X$ such that the evaluation φ_x at x has a unique representing measure. Then $c(A)$ is a boundary, whose closure is ∂A . If X is metrizable, $c(A)$ is a G_δ set in X and supports a representing measure for every member of $\mathfrak{M}(A)$.

For $\varphi \in \mathfrak{M}(A)$, M_φ denotes the set of representing measures for φ . It is a * weak* compact convex subset of the space of measures on X . M_φ is a singleton if A is Dirichlet or logmodular. It is finite-dimensional if A is hypo-Dirichlet. The case $\dim M_\varphi < +\infty$ has been studied in detail. Extensive studies for the case $\dim M_\varphi = +\infty$ have been done only for concrete examples related to polydisks, infinitely connected domains, etc.

The notion of boundary reflects the * maxi-

mum modulus principle for analytic functions, and representing measures come from * Poisson's integral formula. The most relevant maximum principle is H. Rossi's **local maximum modulus principle**: For any closed set K in $\mathfrak{M}(A)$, we have $\|\hat{f}\|_K = \|\hat{f}\|_{\text{bd } K \cap (K \cap \partial A)}$ for all $f \in A$, where $\text{bd } K$ is the topological boundary of K and $\|\hat{f}\|_S = \sup\{|\hat{f}(s)| \mid s \in S\}$. A corresponding result for function spaces was considered by Y. Hirashita and J. Wada. Closely related to representing measures are the **orthogonal measures** for A , i.e., complex Borel measures μ on X such that $\int f d\mu = 0$ for all $f \in A$; they are often useful in studying function algebras by means of the duality technique. The set of orthogonal measures is denoted by A^\perp .

D. Peak Sets [1–3]

A subset K of X is a **peak set** for A if there exists $f \in A$ such that $f(x) = 1$ for $x \in K$ and $|f(y)| < 1$ for $y \in X - K$. K is a **generalized peak set** if it is the intersection of peak sets. A point $x \in X$ is a **(generalized) peak point** if the set $\{x\}$ is a (generalized) peak set. The set of generalized peak points equals the Choquet boundary for A . If X is metrizable, the peak sets and generalized peak sets coincide. There exist X and A such that X is metrizable: $X = \mathfrak{M}(A) = c(A)$, but $A \neq C(X)$ (B. Cole). A subset E of X is **interpolating** for A if for any bounded continuous function u on E there exists $f \in A$ with $f|_E = u$. Then a closed G_δ set K in X is an interpolating peak set if and only if $\mu \in A^\perp$ implies $|\mu|(K) = 0$ (E. Bishop).

E. Antisymmetric Decomposition [1–3]

A subset F of X is a **set of antisymmetry** for A if every function in A which is real-valued on F is constant on F . Bishop's antisymmetric decomposition then appears as an extension of the * Weierstrass-Stone theorem on uniform approximation. It is a refinement of Shilov's decomposition and reads as follows: If $\{E_\alpha\}$ is the family of maximal sets of antisymmetry for A , it is a partition of X into generalized peak sets such that $f \in C(X)$ with $f|_{E_\alpha} \in A|_{E_\alpha}$ for all α belongs to A . An interesting connection was found by J. Tomiyama between the maximal antisymmetric decomposition of X relative to A and that of $\mathfrak{M}(A)$ relative to \hat{A} .

F. Parts and Analytic Structure [1–4]

A. M. Gleason defined an equivalence relation \sim in $\mathfrak{M}(A)$ by setting $\varphi \sim \psi$ if $\sup\{|\hat{f}(\varphi) - \hat{f}(\psi)| \mid f \in A, \|f\| < 1\} < 2$. Each equivalence class for this relation is a **part** (or a **Gleason**

part) for A . Two points φ, ψ belong to the same part if and only if there exist mutually absolutely continuous representing measures μ for φ and ν for ψ such that $c^{-1} < d\mu/d\nu < c$ for some constant $c > 0$ (Bishop). An **analytic structure** in $\mathfrak{M}(A)$ is a pair (V, τ) , with an \dagger analytic set V in some open subset of \mathbb{C}^n and a nonconstant continuous mapping $\tau: V \rightarrow \mathfrak{M}(A)$ such that $\hat{f} \circ \tau$ is analytic on V for all $f \in A$. For such a structure, $\tau(V)$ is always within a part. But there can be a nontrivial part with no analytic structure, as was shown by G. Stolzenberg. A topological characterization of parts was obtained by J. Garnett. Sample results in the positive direction are: (1) If the ideal $A_\varphi = \{f \in A \mid \hat{f}(\varphi) = 0\}$ is finitely generated, there is an analytic structure (V, τ) such that τ is a homeomorphism of V onto an open neighborhood of φ (Gleason); (2) if φ has a unique representing measure and if the part P of φ is not a singleton, there is a bijective continuous mapping τ of the open unit disk onto P such that $\hat{f} \circ \tau$ is analytic for $f \in A$ (J. Wermer, K. Hoffman and G. Lumer); (3) if A is hypo-Dirichlet and if a part P is not a singleton, P can be made into a 1-dimensional \dagger analytic space so that each $f \in A$ is analytic on P (J. Wermer and B. V. O'Neill). Analytic structures in \dagger polynomially convex hulls of curves in \mathbb{C}^n have also been studied [3, 4].

G. Abstract Function Theory [1, 6–9]

Let $\varphi \in \mathfrak{M}(A)$, and choose $m \in M_\varphi$, which is fixed. The **generalized Hardy class** $H_p(m)$, $0 < p \leq \infty$, associated with A is the closure (weak* closure, if $p = \infty$) of A in the $\dagger L_p$ space $L_p(m)$ on the measure space (X, m) . Under suitable restrictions on A , φ , or m , we can recapture some of the important classical facts, most of which have their origins in the works of A. Beurling, R. Nevanlinna, F. and M. Riesz, and G. Szegő. In this area, H. Helson and D. Lowdenslager came up with a powerful method using orthogonal projections in Hilbert space and gave together with S. Bochner's remark, a strong influence for subsequent development. The modification argument was then devised by Hoffman and Wermer, inspired by F. Forelli. After Hoffman's detailed study of logmodular algebras, Lumer observed that most results remain valid when φ has a unique representing measure. T. P. Srinivasan and J.-K. Wang (\rightarrow [8]) introduced the notion of weak* Dirichlet algebra and showed that some major theorems are mutually equivalent and are in fact measure-theoretic. With the Hoffman-Rossi complement, their result now states that for fixed $m \in M_\varphi$ the following are equivalent: (i) $A + \bar{A}$ is weak* dense in

$L_\infty(m)$, i.e., A is a **weak* Dirichlet algebra** on (X, m) ; (ii) if $\mu \in M_\varphi$ is absolutely continuous with respect to m , then $\mu = m$; (iii) if a closed subspace M of $L_2(m)$ is simply invariant in the sense that $A_\varphi M \subseteq M$ and $A_\varphi M$ is not dense in M , then $M = qH_2(m)$ with $q \in L_\infty(m)$, $|q| = 1$ a.e.; (iv) the set $\log|H_\infty(m)^{-1}|$ coincides with $L_\infty(m; \mathbb{R})$, the set of real-valued elements in $L_\infty(m)$; (v) for every $w \in L_1(m)$, $w \geq 0$, we have $\inf\{\int |1 - f|^2 w dm \mid f \in A_\varphi\} = \exp(\int \log w dm)$; (vi) the linear functional $h \rightarrow \int h dm$ on $H_\infty(m)$ has a unique positive extension to $L_\infty(m)$. Further extensions were subsequently made by use of the conjugation operator (\rightarrow Section K) by T. Gamelin, H. König, Lumer, K. Yabuta, and others. Some more properties of weak* Dirichlet algebras have been obtained by T. Nakazi.

H. Generalized Analytic Functions [1, 10]

Let Γ be a dense subgroup of the additive group \mathbb{R} of the reals with discrete topology, and let G be the \dagger character group of Γ . Each $a \in \Gamma$, as a character of G , defines a continuous function χ_a on G . Let A be the closed subalgebra of $C(G)$ generated by $\{\chi_a \mid a \in \Gamma, a \geq 0\}$. A is a Dirichlet algebra but is far more difficult to describe than the disk algebra. The study of this algebra, especially that of invariant subspaces, has evolved from papers by Helson and D. Lowdenslager. Let σ be the \dagger Haar measure of G and $H_2(\sigma)$ the closure of A in $L_2(\sigma)$. A closed subspace M of $L_2(\sigma)$ is called **invariant** if $\chi_a M \subseteq M$ for all $a \in \Gamma$, $a \geq 0$. M is called **doubly invariant** if $\chi_a M \subseteq M$ for all $a \in \Gamma$. Otherwise, it is called **simply invariant**. In fact, only the latter is interesting. Let $e_t \in G$ with $t \in \mathbb{R}$ be the character of Γ defined by $e_t(a) = e^{ita}$. Then the mapping $t \rightarrow e_t$ is a faithful representation of \mathbb{R} into G . A **cocycle** on G is defined to be a Borel function B on $G \times \mathbb{R}$ such that (i) $|B(x, t)| = 1$, (ii) $B(x + e_t, t) = \bar{B}(x, s)B(x, s + t)$ for $x \in G$ and $s, t \in \mathbb{R}$. Two cocycles are identified if they differ only on a null set in $G \times \mathbb{R}$. For a cocycle B , let M_B be the set of $f \in L_2(\sigma)$ such that $\overline{B(x, t)f(x + e_t)} \in H_2(dt/(1 + t^2))$ for almost all $x \in G$, where $H_2(dt/(1 + t^2))$ is the closure, in the space $L_2(dt/(1 + t^2))$ on \mathbb{R} , of the set of boundary value functions on \mathbb{R} of bounded analytic functions on the upper half-plane. Then the mapping $B \rightarrow M_B$ is a bijection from the set of cocycles onto the set of simply invariant subspaces M of $L_2(\sigma)$ such that $M = \bigcap \{\chi_a M \mid a \in \Gamma, a < 0\}$. Moreover, $M_B = qH_2(\sigma)$ for some $q \in L_\infty(\sigma)$, $|q| = 1$ a.e. if and only if $B(x, t) = \overline{q(x)}q(x + e_t)$, i.e., B is a **coboundary**. When $\Gamma \neq \mathbb{R}$, there is a cocycle that is not a coboundary. Further studies have been done by Helson,

Gamelin, J. Tanaka, and others. On the other hand, F. Forelli observed that flows in compact spaces give rise to a kind of analyticity. In the above case, the algebra A consists of all $f \in C(G)$ that are analytic with respect to the flow $T_t(x) = x + e_t$ for $x \in G$ and $t \in \mathbb{R}$. Generalized analytic functions induced by flows in general have been studied by Forelli and P. S. Muhly.

I. The Unit Disk [1, 6, 7, 11, 12]

Let A be the disk algebra $P(\mathbb{D})$, \mathbb{D} the open unit disk $\{|z| < 1\}$, and m the normalized Lebesgue measure on \mathbb{T} . The algebra A has been the most important model in the theory of function algebras, and we find here the origin of many abstract results. Some typical results: (i) every orthogonal measure for A is absolutely continuous with respect to m (F. and M. Riesz); (ii) a closed set E in \mathbb{T} is an interpolating peak set for the Gel'fand transform \hat{A} of A if and only if $m(E) = 0$ (W. Rudin and L. Carleson); (iii) A is maximal among the closed subalgebras of $C(\mathbb{T})$ (Wermer); (iv) a function $f \in C(\text{Cl } \mathbb{D})$ belongs to \hat{A} if f is analytic at every $z \in \mathbb{D}$ with $f(z) \neq 0$ (T. Radó).

The generalized Hardy class $H_p(m)$, $0 < p \leq \infty$, associated with A is viewed as the set of nontangential boundary value functions of elements in the classical Hardy class $H_p(\mathbb{D})$. Here we find the origin of invariant subspace theorems: A closed subspace $M (\neq \{0\})$ of $H_2(m)$ is invariant, i.e., $AM \subseteq M$, if and only if $M = qH_2(m)$ with $q \in H_\infty(m)$, $|q| = 1$ a.e. (Beurling).

The algebra $H_\infty = H_\infty(m)$ is a weak* Dirichlet algebra, whose Shilov boundary is identified with the maximal ideal space X of $L_\infty(m)$. We have $L_\infty(m; \mathbb{R}) = \log |(H^\infty)^{-1}|$, and a fortiori H_∞ is logmodular on X . The mapping $z \rightarrow \varphi_z$ embeds the disk \mathbb{D} in $\mathfrak{M}(H_\infty)$ as an open set. The structure of $\mathfrak{M}(H_\infty)$ was studied in detail by I. J. Schark, Hoffman, and others. We finish with three remarkable results: (i) \mathbb{D} is dense in $\mathfrak{M}(H_\infty)$ (Carleson). This is the **corona theorem** and was proved in the following equivalent form: For any $f_1, \dots, f_n \in H_\infty(\mathbb{D})$ with $|f_1| + \dots + |f_n| \geq \varepsilon > 0$ on \mathbb{D} , there exist $g_1, \dots, g_n \in H_\infty(\mathbb{D})$ with $f_1 g_1 + \dots + f_n g_n = 1$. A simple proof was discovered by T. Wolff [12]. (ii) The convex combinations of Blaschke products are uniformly dense in the unit ball of $H_\infty(\mathbb{D})$ (D. Marshall) [12]. (iii) Every closed subalgebra B between H_∞ and $L_\infty(m)$ is a **Douglas algebra**, i.e., B is generated by H_∞ and the complex conjugates of a family of inner functions (S.-Y. Chang and Marshall) [11]. The proof of (iii) is an interesting application of the theory of bounded mean oscillation.

J. Rational Approximation [1, 2]

The problem of rational (polynomial) approximation on a compact plane set K asks when $A(K) = R(K)$ ($A(K) = P(K)$) holds. An important tool for such problems is **Cauchy's transform** of measure μ : $\hat{\mu}(\zeta) = \int (z - \zeta)^{-1} d\mu(z)$. Using this, one can show, for instance, the following: (1) $A(K) = P(K)$ if and only if K has a connected complement (S. N. Mergelyan); (2) $f \in C(K)$ belongs to $R(K)$ if each point $z \in K$ has a closed neighborhood V with $f|_{K \cap V} \in R(K \cap V)$ (Bishop); (3) $A(K) = R(K)$ if the diameters of the components of $\mathbb{C} - K$ are bounded away from zero (Mergelyan); (4) $C(K) = R(K)$ if and only if almost all points of K are peak points for $R(K)$ (Bishop). The last result cannot be extended much because there is a set K such that $A(K) \neq R(K)$, while $A(K)$ and $R(K)$ have the same peak points (A. M. Davie).

A complete characterization for $A(K) = R(K)$ was obtained by A. G. Vitushkin: (5) The following are equivalent: (i) $A(K) = R(K)$; (ii) for any bounded open set D in \mathbb{C} , $\alpha(D - K) = \alpha(D - K^0)$; (iii) for any $z \in \text{bd } K$, there exists $r \geq 1$ such that $\limsup_{\delta \rightarrow 0} \alpha(\Delta(z; \delta) - K^0) / \alpha(\Delta(z; r\delta) - K) < +\infty$, where $\Delta(z; \delta)$ is the disk $\{w \in \mathbb{C} \mid |w - z| \leq \delta\}$ and $\alpha(E)$, for any bounded set E in \mathbb{C} , is the **continuous analytic capacity** of E , which is the supremum of $|f'(\infty)|$ for all continuous functions f on the Riemann sphere $\mathbb{C} \cup \{\infty\}$ such that $|f| \leq 1$, $f(\infty) = 0$, and f is analytic off a compact subset of E . As for uniform or asymptotic approximation on noncompact closed subsets in \mathbb{C} , Carleman's classical study has recently been extended by N. U. Arakelyan, A. A. Nersesyan, A. Stray, and others in an interesting way.

In connection with rational approximation, we should note detailed studies on pointwise bounded approximation in $H_\infty(U)$, U being a bounded open set in \mathbb{C} , by O. J. Farrell, L. Rubel and A. Shields, Gamelin and J. Garnett, and Davie.

K. Further Topics

(1) **Conjugation operator** [9, 13]. Take any $m \in M_\varphi$, $\varphi \in \mathfrak{M}(A)$. The **conjugation operator** then associates with each $u \in \text{Re } A$ the unique element $*u$ in $C_R(X)$ such that $u + i*u \in A$ and $\int *u dm = 0$. After the classical inequalities of Kolmogorov and M. Riesz, we consider the following conditions with constants c_p and d_p : (K) $(\int |*u|^p dm)^{1/p} \leq c_p \int |u| dm$ for $0 < p < 1$; (M) $(\int |*u|^p dm)^{1/p} \leq d_p (\int |u|^p dm)^{1/p}$ for $1 < p < \infty$. Let m be arbitrary. Then the inequality (K) is valid for $u \in \text{Re } A$, $u > 0$. The inequality (M) is valid for all $u \in \text{Re } A$ if p is an even integer ≥ 2 ; it is

valid for all $u \in \operatorname{Re} A$, $u > 0$, if p is not an odd integer. All remaining cases have counterexamples. On the other hand, M_φ always contains an m such that $\log|f(\varphi)| \leq \int \log|f| dm$ for $f \in A$ (Bishop). Such a representing measure is called a **Jensen measure** for φ . If m is Jensen, (K) is valid for all $u \in \operatorname{Re} A$ and all $0 < p < 1$, and (M) is valid for all $u \in \operatorname{Re} A$ and all $1 < p < \infty$.

(2) **Riemann surfaces.** For a compact bordered Riemann surface R , let $A(R)$ be the algebra of functions, continuous on $\operatorname{Cl} R$ and analytic on R . Then $A(R)$ is hypo-Dirichlet on the border $\operatorname{bd} R$ of R , and many results for the disk algebra are extended to $A(R)$. Most of the basic results are described in [14]. The maximality of $A(R)$ in $C(\operatorname{bd} R)$ was obtained by H. Royden; extreme points of related Hardy classes were discussed by Gamelin and M. Voichick; and invariant subspaces were determined by Forelli, M. Hasumi, D. Sarason, and Voichick. A further extension to infinitely connected surfaces has been obtained by C. W. Neville, Hasumi, and M. Hayashi in the case of open Riemann surfaces R of Parreau-Widom type, which is defined as follows: Let $G(a, z)$ be \dagger Green's function for R , and let $B(a, \alpha)$, $\alpha > 0$, be the first \dagger Betti number of the domain $\{z \in R \mid G(a, z) > \alpha\}$; then R is of **Parreau-Widom type** if $\int B(a, \alpha) d\alpha < +\infty$. For such surfaces the situation looks favorable: For instance, the Cauchy-Read theorem is valid, and the Brelot-Choquet problem concerning Green's lines is solved affirmatively.

As for the generalization of approximation theorems of Mergelyan and Arakelyan, we refer to the work of Bishop and L. K. Kodama for compact sets and to that of S. Scheinberg for noncompact closed sets.

(3) **Higher-dimensional sets.** Much attention has been paid to algebras of analytic functions on domains in \mathbb{C}^n , $n \geq 2$, e.g., polydisks, unit balls, and general pseudoconvex domains. Polydisk algebras and ball algebras have been studied extensively by W. Rudin, P. Ahern, Forelli, and many others [15, 16]. Approximation theorems of Mergelyan type were obtained by G. M. Henkin, N. Kerzman, and I. Lieb for strictly pseudoconvex domains with smooth boundary and by L. Hörmander and Wermer and L. Nirenberg and R. O. Wells, Jr., in the case of totally real manifolds. Further improvements have been obtained by R. M. Range, A. Sakai, and others.

References

[1] T. W. Gamelin, Uniform algebras, Prentice-Hall, 1969.

- [2] A. Browder, Introduction to function algebras, Benjamin, 1969.
- [3] E. L. Stout, The theory of uniform algebras, Bogden and Quigley, 1971.
- [4] J. Wermer, Banach algebras and several complex variables, Markham, 1971.
- [5] T. W. Gamelin, Lectures on $H^\infty(D)$, Universidad Nacional de La Plata, 1972.
- [6] K. Hoffman, Banach spaces of analytic functions, Prentice-Hall, 1962.
- [7] H. Helson, Lectures on invariant subspaces, Academic Press, 1964.
- [8] F. Birtel (ed.), Function algebras (Proc. Int. Symposium on Function Algebras, Tulane Univ., 1965), Scott, Foresman, 1966.
- [9] K. Barbey and H. König, Abstract analytic function theory and Hardy algebras, Lecture notes in math. 593, Springer, 1977.
- [10] H. Helson, Analyticity on compact Abelian groups, Algebras in Analysis, J. H. Williamson (ed.), Academic Press, 1975, 1–62.
- [11] D. Sarason, Function theory on the unit circle, Virginia Polytech. Inst. and State Univ., 1978.
- [12] P. Koosis, Introduction to H_p spaces, Cambridge Univ. Press, 1980.
- [13] T. W. Gamelin, Uniform algebras and Jensen measures, Cambridge Univ. Press, 1978.
- [14] M. Heins, Hardy classes on Riemann surfaces, Lecture notes in math. 98, Springer, 1969.
- [15] W. Rudin, Function theory in polydisks, Benjamin, 1969.
- [16] W. Rudin, Function theory in the unit ball of \mathbb{C}^n , Springer, 1980.

165 (II.4) Functions

A. History

Leibniz used the term *function* (Lat. *functio*) in the 1670s to refer to certain line segments whose lengths depend on lines related to curves. Soon the term was used to refer to dependent quantities or expressions. In 1718, Johann Bernoulli used the notation φx , and by 1734 the modern functional notation $f(x)$ had been used by Clairaut and by Euler, who defined functions as analytic formulas constructed from variables and constants (1728) [1]. \dagger Cauchy stated (1821) [2]: "When there is a relation among many variables, which determines along with values of one of them the values of the others, we usually consider the others as expressed by the one. We then call

the one an 'independent variable,' and the others 'dependent variables.'" †Dirichlet considered a function of $x \in [a, b]$ in his paper (1837) [3] concerning representations of "completely arbitrary functions" and stated that there was no need for the relation between y and x to be given by the same law throughout an interval, nor was it necessary that the relation be given by mathematical formulas. A function was simply a correspondence in which values of one variable determined values of another.

B. Functions

Today, the word "function" is used generally in mathematics in the same sense as a †mapping (\rightarrow 381 Sets C) or, which is the same thing, a †univalent correspondence (\rightarrow 358 Relations B). But this word is sometimes used in a wider sense, to mean a general (not necessarily univalent) correspondence, called a **many-valued** (or **multivalued**) **function**; in that case a univalent correspondence is called a **single-valued function**.

Specialists in each branch of mathematics have their respective ways of using the word. In analysis, values of a function are often considered real or complex numbers; such functions are called **real-valued functions** or **complex-valued functions**, respectively. Furthermore, if the domain of the function is also a set of real or complex numbers, then it is called a **real function** or a **complex function**, respectively (\rightarrow 131 Elementary Functions; 84 Continuous Functions; 198 Holomorphic Functions). If the domain of a real- or complex-valued function is contained in a †function space, the function is often called a **functional**; the †distribution is an example. In algebra we often fix a †field, †ring, etc., and consider functions whose domains and ranges are in such algebraic systems. Special names are given to functions having special properties, which can be defined according to the structures of the domain and the range. For example, when both domain and range of a function f are sets of real numbers, f is called an **even function** if $f(t) = f(-t)$, and an **odd function** if $f(t) = -f(-t)$. A function f that preserves the order relation between real numbers, i.e., such that $t_1 < t_2$ implies $f(t_1) \leq f(t_2)$, is called a †monotone increasing function.

A mapping from a set I to a set F of functions, $\varphi: I \rightarrow F$, is called a **family of functions indexed by I** (or simple **family of functions**), and is denoted, using the form f_λ instead of $\varphi(\lambda)$, by $\{f_\lambda\}_{\lambda \in I}$ or $\{f_\lambda\}_{\lambda \in I}$. In particular, if I is the set of natural numbers, the family is called a **sequence of functions**.

C. Variables

A letter x , for which we can substitute a name of an element of a set X , is called a **variable**, and X is called the **domain** of the variable. An element of the domain of a variable x is called a **value** of x . In particular, if the domain is a set of real numbers or complex numbers, the variable is called a **real variable** or a **complex variable**, respectively. On the other hand, a letter that stands for a particular element is called a **constant**.

When the domain and range of a function f are X and Y , respectively, a variable x whose domain is X is called the **independent variable**, and a variable y whose domain is Y is called the **dependent variable**. Then we say y is a function of x , and write $y = f(x)$. When a concrete method is given by which we make a value of y correspond to each value of x , we say that y is an **explicit function** of x . When a function is determined only by a †binary relation such as $R(x, y) = 0$, we say that y is an **implicit function** of x (\rightarrow 208 Implicit Functions).

Given functions f, g with an independent variable t , suppose that y is regarded as a function of x defined by relations $x = f(t)$, $y = g(t)$. Then we say that y is a function of x with the variable t as a **parameter**. A function whose range is a given set C with variable t as its independent variable is often called a **parametric representation** of C by t .

If the domain of a function f is contained in a Cartesian product set $X_1 \times X_2 \times \dots \times X_n$, the independent variable is denoted by (x_1, x_2, \dots, x_n) , and f is often called a **function of n variables** or a **function of many variables** (when $n \geq 2$).

D. Families and Sequences

A function whose domain is a set I , $\varphi: I \rightarrow X$, is called a **family indexed by I** (or simply **family**), and I is called the **index set**. In the case $\varphi(\lambda) = x_\lambda (\lambda \in I)$, the family is denoted by $\{x_\lambda\}_{\lambda \in I}$ or $\{x_\lambda\}_{\lambda \in I}$. If the range X of a function φ is a set of points, a set of functions, a set of mappings, or a set of sets, then the family $\{x_\lambda\}_{\lambda \in I}$ is called a **family of points**, a **family of functions**, a **family of mappings**, or a **family of sets**, respectively. If the set I is a †directed set, the family is called a **directed family**. Generally, if J is a subset of I , the family $\{x_\lambda\}_{\lambda \in J}$ is called a **subfamily** of $\{x_\lambda\}_{\lambda \in I}$. In particular, if I is a finite or infinite set of natural numbers, the family indexed by I is called a **finite sequence** or **infinite sequence**, respectively. **Sequence** is a generic name for both, but in many cases it means an infinite sequence, and usually we

have $I = \mathbf{N}$. Then the value corresponding to $n \in \mathbf{N}$ is called the **n th term** or, generally, a **term**. For convenience, the 0th term is often used as well. If each term of a sequence is a number, a point, a function, or a set, the sequence is called a **sequence of numbers**, a **sequence of points**, a **sequence of functions**, or a **sequence of sets**, respectively. A sequence is usually denoted by $\{a_n\}$. If it is necessary to show the domain of n explicitly, the sequence is denoted by $\{a_n\}_{n \in I}$. If J is a subset of I , a sequence $\{a_n\}_{n \in J}$ is called a **subsequence** of the sequence $\{a_n\}_{n \in I}$. And if $I = \mathbf{N}$, the composite $\{a_{k_n}\}$ of $\{a_n\}$ and a sequence $\{k_n\}$ of natural numbers with $k_1 < k_2 < k_3 \dots$ is usually called a subsequence of $\{a_n\}$.

References

- [1] L. Euler, Opera omnia, ser. I, Opera mathematica VIII: Introductio in analysin infinitorum I, Teubner, 1922.
 [2] A. L. Cauchy, Cours d'analyse de l'Ecole Royale Polytechnique pt. 1, 1821 (Oeuvres ser. 2, III, Gauthier-Villars, 1897).
 [3] G. P. L. Dirichlet, Über die Darstellung ganz willkürlicher Funktion durch Sinus- und Cosinusreihen, Repertorium der Physik, Bd. 1 (1837), 152–174 (Werke, Bd. 1, G. Reimer, 1889, I, p. 133–160).

166 (X.5) Functions of Bounded Variation

A. Monotone Functions

A function (or mapping) f from an \dagger ordered set X to another ordered set Y is called a **monotone increasing (monotone decreasing) function** if

$$\begin{aligned} x_1 < x_2 & \text{ implies } f(x_1) \leq f(x_2) \\ (x_1 < x_2 & \text{ implies } f(x_1) \geq f(x_2)). \end{aligned} \quad (1)$$

A monotone increasing (decreasing) function is also called a **nondecreasing (nonincreasing) function**. In either case, the function f is called simply a **monotone function**. If X and Y are \dagger totally ordered sets and the inequality $< (>)$ holds in (1) instead of $\leq (\geq)$, then f is called a **strictly (monotone) increasing (strictly (monotone) decreasing) function**. In either case, f is called simply a **strictly monotone function**.

In particular, when X and Y are subsets of the real line \mathbf{R} , a monotone function is contin-

uous except for at most a countable number of points. Hence it is \dagger Riemann integrable in a finite interval provided that it is bounded. A continuous real function $f(x)$ defined on an interval in \mathbf{R} is \dagger injective if and only if it is strictly monotone. In such a case, the range of the function $f(x)$ is also an interval, and the inverse function is also strictly monotone. Furthermore, a differentiable real function f defined on an interval is monotone if and only if its derivative f' is always ≥ 0 (monotone increasing) or always ≤ 0 (monotone decreasing). If $f' > 0$ (< 0), f is strictly monotone increasing (decreasing).

B. Functions of Bounded Variation

Let $f(x)$ be a real function defined on a closed interval $[a, b]$ in \mathbf{R} . Given a subdivision of the interval $a = x_0 < x_1 < x_2 < \dots < x_n = b$, we denote the sum of positive differences $f(x_i) - f(x_{i-1})$ by P and the sum of negative differences $f(x_i) - f(x_{i-1})$ by $-N$. Then we easily obtain

$$P - N = f(b) - f(a),$$

$$P + N = \sum_i |f(x_i) - f(x_{i-1})|.$$

The suprema of P , N , and $P + N$ for all possible subdivisions of $[a, b]$ are called the **positive variation**, the **negative variation**, and the **total variation** of the function $f(x)$ in the interval $[a, b]$, respectively. If any of these three values is finite, then all three values are finite. In such a case, the function $f(x)$ is called a **function of bounded variation**. Every function of bounded variation is bounded, but the converse is not true. The positive and negative variations $\pi(t)$, $v(t)$ of the function $f(x)$ in the interval $[a, t]$ are monotone increasing functions with respect to t , and we have

$$f(x) - f(a) = \pi(x) - v(x) \quad (2)$$

if $f(x)$ is a function of bounded variation. Hence every function of bounded variation has both left and right limits at every point. A monotone function is a function of bounded variation, and the sum, the difference, or the product of two functions of bounded variation is also a function of bounded variation. Hence $f(x)$ is a function of bounded variation if and only if it is the difference of two monotone functions. The representation (2) (representing a function of bounded variation as the difference of two monotone increasing functions) is called the **Jordan decomposition** of the function $f(x)$. A function of bounded variation is Riemann integrable, continuous except for at

most a countable number of points, and differentiable †almost everywhere.

A continuous function defined on a closed interval is bounded but not necessarily of bounded variation (e.g., $f(x) = x \sin(1/x)$ ($x \in (0, 1]$, 0 ($x = 0$)). A discontinuous function may be a function of bounded variation on a closed interval (e.g., $\text{sgn}(x)$). However, an †absolutely continuous function, a differentiable function with bounded derivative, or a function satisfying the †Lipschitz condition is a function of bounded variation on a closed interval.

The notion of functions of bounded variation was introduced by C. Jordan in connection with the notion of the length of curves (\rightarrow 246 Length and Area).

C. Lebesgue-Stieltjes Integral

Let $f(x)$ be a right continuous function of bounded variation on a closed interval $[a, b]$, and $f(x) = \pi(x) - v(x)$ the Jordan decomposition of $f(x)$. Then $\pi(x)$ and $v(x)$ are monotone increasing right continuous functions and hence define bounded measures $d\pi(x)$ and $dv(x)$ on $[a, b]$, respectively (\rightarrow 270 Measure Theory L (v)). The difference $d\pi - dv$ of these two measures is a †completely additive set function on $[a, b]$ which is often called the (**signed**) **Lebesgue-Stieltjes measure** induced by f , written df . For every function g integrable with respect to the measure $d\mu = d\pi + dv$, we define $\int_E g df$ to be equal to $\int_E g d\pi - \int_E g dv$ and call it the **Lebesgue-Stieltjes integral**. In this case, $h(x) = \int_{[a, b]} g df$ is of bounded variation on $[a, b]$ and the Lebesgue-Stieltjes measure $dh(x)$ is denoted by $g(x)df(x)$. If f_1 and f_2 are of bounded variation on $[a, b]$, then so is the product $f = f_1 f_2$, and we have

$$df(x) = f_1(x \pm 0)df_2(x) + f_2(x \mp 0)df_1(x). \quad (3)$$

The Lebesgue-Stieltjes integral for a continuous integrand is often called the **Riemann-Stieltjes integral**, because it can be defined in an elementary way similar to the definition of the Riemann integral.

The notion of bounded variation can also be defined for interval functions on \mathbf{R}^n and set functions on an abstract space (\rightarrow 380 Set Functions).

References

- [1] H. L. Royden, Real analysis, Macmillan, second edition, 1963.
- [2] W. Rudin, Real and complex analysis, McGraw-Hill, second edition, 1974.

167 (XIV.7)

Functions of Confluent Type

A. Confluent Hypergeometric Functions

If some singularities of an ordinary differential equation of †Fuchsian type are confluent to each other, we obtain a **confluent differential equation** whose solutions are called **functions of confluent type**. The equations that appear frequently in practical problems are the **confluent hypergeometric differential equations**

$$z \frac{d^2 w}{dz^2} + (\gamma - z) \frac{dw}{dz} - \alpha w = 0 \quad (1)$$

and related equations. Equation (1) corresponds to the †hypergeometric differential equation for which a †regular singular point coincides with the point at infinity and is an †irregular singular point of class 1. For (1), $z = 0$ is a regular singular point, and a series solution (radius of convergence ∞) is given by

$$F(\alpha, \gamma; z) = {}_1F_1(\alpha, \gamma; z) = \sum_{n=0}^{\infty} \frac{\alpha(\alpha+1)\dots(\alpha+n-1)}{n!\gamma(\gamma+1)\dots(\gamma+n-1)} z^n \quad (|z| < \infty), \quad (2)$$

where γ is not a nonpositive integer. The function ${}_1F_1$ in (2) is a †generalized hypergeometric function due to Barnes and is called a **hypergeometric function of confluent type** or **Kummer function**. If γ is not equal to a positive integer, the other solution of (1) independent of (2) is given by $z^{1-\gamma} F(1+\alpha-\gamma, 2-\gamma; z)$ (\rightarrow Appendix A, Table 19.I).

B. Whittaker Functions

Equation (1) with $w = e^{z/2} z^{-\gamma/2} W$, $\gamma - 2\alpha = 2k$, $\gamma^2 - 2\gamma = 4m^2 - 1$ reduces to **Whittaker's differential equation**

$$\frac{d^2 W}{dz^2} + \left(-\frac{1}{4} + \frac{k}{z} + \frac{\frac{1}{4} - m^2}{z^2} \right) W = 0. \quad (3)$$

If $2m$ is not equal to an integer, (3) has two series solutions for any finite z :

$$M_{k,m}(z) = z^{(1/2)+m} e^{-z/2} F\left(\frac{1}{2} + m - k, 1 + 2m; z\right),$$

$$M_{k,-m}(z) = z^{(1/2)-m} e^{-z/2} F\left(\frac{1}{2} - m - k, 1 - 2m; z\right).$$

If $2m$ is an integer, since the functions $M_{k,m}$ and $M_{k,-m}$ are linearly dependent, E. T. Whittaker considered a solution of the form

$$W_{k,m}(z) = -\frac{1}{2\pi i} \Gamma(k + \frac{1}{2} - m) z^k e^{-z/2} \times \int_{\infty}^{(0+)} (-t)^{-k-(1/2)+m} \left(1 + \frac{t}{z}\right)^{k-(1/2)+m} e^{-t} dt.$$

If $k - \frac{1}{2} - m$ is equal to a negative integer, this integral does not exist. The function

$$W_{k,m}(z) = \frac{z^k e^{-z/2}}{\Gamma(\frac{1}{2} - k + m)} \int_0^\infty t^{-k-(1/2)+m} \times \left(1 + \frac{t}{z}\right)^{k-(1/2)+m} e^{-t} dt$$

for $\text{Re}(k - \frac{1}{2} - m) \leq 0$ is defined for any m, k , and for any z except when z is a negative real number. We call $M_{k,m}$ and $W_{k,m}$ the **Whittaker functions**. †Bessel functions are particular cases of these functions, and the relation

$$J_n(z) = \frac{z^{-1/2}}{2^{2n+(1/2)} \Gamma(n+(1/2)) \Gamma(n+1)} M_{0,n}(2iz)$$

is satisfied. In Whittaker's differential equation, since $W_{-k,m}(-z)$ is also a solution and $W_{k,m}(z)/W_{-k,m}(-z)$ is not equal to a constant, $W_{k,m}(z)$ and $W_{-k,m}(-z)$ can be considered a pair of fundamental solutions (\rightarrow Appendix A, Table 19.II).

C. Parabolic Cylinder Functions

Putting $x = (\xi^2 - \eta^2)/2$ and $y = \xi\eta$, the curves corresponding to $\xi = \text{constant}$ and to $\eta = \text{constant}$, respectively, constitute families of orthogonal parabolas. The curvilinear coordinates (ξ, η, z) in three dimensions are called **parabolic cylindrical coordinates**. By using parabolic coordinates, separating variables in Laplace's equation into the form $f(\xi)g(\eta)e^{iz}$, and making a simple transformation, we find that f and g satisfy a differential equation of the form

$$\frac{d^2 F}{dz^2} + (n + \frac{1}{2} - \frac{1}{4}z^2)F = 0. \quad (4)$$

By means of the Whittaker function $W_{k,m}(z)$, a solution $D_n(z)$ of (4) is represented by

$$D_n(z) = 2^{n/2+(1/4)} z^{-1/2} W_{n/2+(1/4), -1/4}(\frac{1}{2}z^2).$$

Equation (4) is called **Weber's differential equation** or the **Weber-Hermite differential equation**, and $D_n(z)$ the **Weber function**. Another solution of (4) is $D_{-n-1}(iz)$ or $D_{-n-1}(-iz)$. The solutions of (4) are called **parabolic cylinder functions**. In particular, if n is equal to a nonnegative integer, then

$$H_n(z) = 2^{-n/2} \exp(\frac{1}{2}z^2) D_n(\sqrt{2}z)$$

is the †Hermite polynomial of degree n . Solutions of differential equations for harmonic oscillators in quantum mechanics are of this form.

In general, suppose that three regular singular points are confluent to the point at infinity, and that they are reduced to an irregular singular point of class 2. Suppose further that

there are no other singularities. Then differential equations of order 2 with these conditions are transformed into the form (4), whose solutions are represented by parabolic cylinder functions. Differential equations of the form (4) are reduced to confluent hypergeometric differential equations if z^2 is chosen as an independent variable (\rightarrow Appendix A, Table 20.III).

D. Indefinite Integrals of Elementary Functions

Since exponential and trigonometric functions can be represented by particular types of Kummer functions, their indefinite integrals that cannot be represented by elementary functions, e.g., †incomplete Γ -functions and the **error function** $\text{Erf } z = \int_0^z \exp(-t^2) dt$, can be represented by Kummer or Whittaker functions. They are included in a family of †special functions of confluent type. The functions defined by

$$C(z) = \frac{1}{\sqrt{2\pi}} \int_0^z \frac{\cos t}{\sqrt{t}} dt,$$

$$S(z) = \frac{1}{\sqrt{2\pi}} \int_0^z \frac{\sin t}{\sqrt{t}} dt$$

are called **Fresnel integrals**, which are also represented in terms of the Whittaker function as

$$C(z) - iS(z) = \frac{1-i}{2} \left(1 - \frac{e^{-\pi i/8}}{\sqrt{\pi}} z^{1/4} e^{-z^2/2} W_{-1/4, 1/4}(iz) \right).$$

Fresnel integrals first appeared in the theory of the diffraction of waves. More recently they have been applied to designing highways for high-speed automobiles. Furthermore, the functions

$$C(u) = \int_0^u \cos\left(\frac{\pi}{2}s^2\right) ds,$$

$$S(u) = \int_0^u \sin\left(\frac{\pi}{2}s^2\right) ds$$

(obtained by a change of variables $z = \pi u^2/2$) are also called Fresnel integrals. Numerical tables are available for them. The curves $x = C$ and $y = S$ with a parameter z or u are called **Cornu's spiral** (Fig. 1). The functions

$$\text{Li } x = \int_0^x \frac{dt}{\log t}, \quad \text{Ei } x = \int_{-\infty}^x \frac{e^t}{t} dt,$$

where a †principal value must be taken at $t = 0$ if $x > 0$,

$$\text{Si } x = \int_0^x \frac{\sin t}{t} dt, \quad \text{and} \quad \text{Ci } x = - \int_x^\infty \frac{\cos t}{t} dt$$

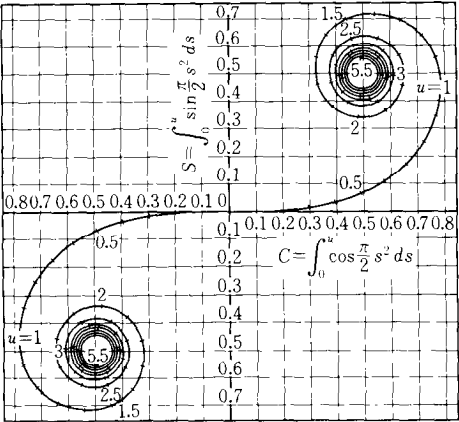


Fig. 1

are called the **logarithmic integral**, **exponential integral**, **sine integral**, and **cosine integral**, or **integral logarithm**, **integral exponent**, **integral sine**, and **integral cosine**, respectively. They satisfy the relations

$Ei\,x = Li\,e^x, \quad Ei\,ix = Ci\,x + i\,Si\,x + (\pi/2)i.$

They have important applications: $Ei\,x$ in quantum mechanics, $Si\,x$ and $Ci\,x$ in electrical engineering, and $Li\,x$ in estimating the number of †primes less than x (→ 123 Distribution of Prime Numbers). $Li\,x$ is also denoted by $li\,x$ (→ Appendix A, Table 19.II).

E. Stokes's Equation

Consider a linear differential equation of the second order with five regular singular points including the point at infinity such that the difference of the characteristic indices at every singularity is equal to 1/2. Such equations are called **generalized Lamé's differential equations**. F. Klein and M. Bôcher have shown that every linear differential equation that is commonly treated in mathematical physics is represented by a confluent type of generalized Lamé's equation. Among these equations, if all five singularities are confluent to the point at infinity, the resulting equation is called **Stokes's differential equation**, which is applied to the investigation of diffraction. This is reduced to †Bessel's differential equation of order 1/3 by suitable transformations of the independent and dependent variables.

References

[1] H. Buchholz, The confluent hypergeometric function, with special emphasis on its applications, translated by H. Lichtblau and K. Wentzel, Springer, 1969. (Original in German, 1953.)

[2] L. J. Slater, Confluent hypergeometric functions, Cambridge Univ. Press, 1960. For the logarithmic integral, etc.,
[3] N. Nielsen, Theorie der Integrallogarithmus und verwandter Transzendenten, Teubner, 1906.
[4] Mathematical tables I, British Assoc. Adv. Sci., London, 1931.
[5] Nat. Bur. Standards, Tables of sine, cosine and exponential integrals I, II, Washington, D. C., 1940.
[6] L. Lewin, Polylogarithms and associated functions, North-Holland, 1958, revised edition, 1981.
Also → references to 389 Special Functions.

168 (XII.6)
Function Spaces

A. General Remarks

It is a general method in modern analysis to consider a set X of mappings of a space Ω into another space A as a space (→ 381 Sets) and its elements (namely, mappings of Ω into A) as points of the space X , and to investigate them as geometric objects. In particular, it is important to consider the case where Ω is a †topological space, a †measure space, or a †differentiable manifold and X is a set of real- or complex-valued functions defined on Ω and satisfying certain conditions, such as continuity, measurability, and differentiability. Such spaces are generally called **function spaces**; they usually form †topological linear spaces (→ 37 Banach Spaces; 197 Hilbert Spaces; 424 Topological Linear Spaces).

B. Examples of Function Spaces

The following are important examples of function spaces. Throughout this section, all functions are real- or complex-valued, and two functions on a measure space are identified whenever they are equal to each other †almost everywhere.

(1) **The Function Spaces $C(\Omega)$, $C_\infty(\Omega)$, and $C_0(\Omega)$.** The totality of continuous functions $f(x)$ defined on a compact †Hausdorff space Ω is denoted by $C(\Omega)$. Let $f + g$ and αf (α a real or complex number) be the functions $f(x) + g(x)$ and $\alpha f(x)$, respectively. Then $C(\Omega)$ forms a †linear space. Furthermore, define the norm of f by $\|f\| = \sup_{x \in \Omega} |f(x)|$. Then $C(\Omega)$ becomes a †Banach space since it is complete in (the

metric defined by) the norm. The norm is called the **supremum norm** or **uniform norm** because $\lim_{n \rightarrow \infty} \|f_n - f\| = 0$ means that $f_n(x)$ converges to $f(x)$ uniformly on Ω as $n \rightarrow \infty$. Define $f \cdot g$ to be the function $f(x)g(x)$. Then clearly $\|f \cdot g\| \leq \|f\| \|g\|$. Hence $C(\Omega)$ is also a \dagger Banach algebra.

Suppose that a subset R of $C(\Omega)$ satisfies the following three conditions: (i) R is a \dagger algebra over the complex number field with respect to the addition and multiplication defined above and contains the function identically equal to one. (ii) For any two distinct points x and y of Ω , there exists a function $f \in R$ satisfying $f(x) \neq f(y)$. (iii) For any $f \in R$, there exists an $f^* \in R$ such that $f^*(x) = \overline{f(x)}$ on Ω . Then R is dense in $C(\Omega)$ with respect to the supremum norm (namely, in the sense of uniform convergence). This fact is known as the **Weierstrass-Stone theorem** (or **Stone-Gel'fand theorem**). A subset E of $C(\Omega)$ is \dagger precompact (i.e., any sequence of functions in E contains a subsequence that converges uniformly on Ω) if and only if E is \dagger uniformly bounded and \dagger equicontinuous (**Ascoli-Arzelà theorem**). Since $C(\Omega)$ is not a \dagger reflexive Banach space except for trivial cases (\rightarrow Section C), precompact sets and relatively compact sets are different in the \dagger weak topology. For important characterizations of the latter sets \rightarrow [2, 5].

When Ω is a topological space that is not necessarily compact, the totality of bounded continuous functions on Ω (denoted by $BC(\Omega)$) is also a Banach space with respect to the supremum norm $\|f\| = \sup_{x \in \Omega} |f(x)|$. Let Ω be a locally compact Hausdorff space. Then the space $C(\Omega)$ of all continuous functions on Ω is endowed with the topology of uniform convergence on the compact sets, i.e., the \dagger locally convex topology defined by the \dagger seminorms $\sup_{x \in K} |f(x)|$ as K ranges over the compact sets in Ω . $C(\Omega)$ is always a complete locally convex space. It is a \dagger Fréchet space if Ω is σ -compact (i.e., Ω is the union of a countable family of compact sets). We denote by $C_\infty(\Omega)$ the subspace of all functions $f(x) \in C(\Omega)$ that converge to zero as x tends to infinity (i.e., given an $\varepsilon > 0$, there is a compact set K such that $|f(x)| < \varepsilon$ for $x \notin K$). $C_\infty(\Omega)$ is a Banach space with the norm $\sup_{x \in \Omega} |f(x)|$. It can be regarded as a closed linear subspace of $BC(\Omega)$.

The totality of continuous functions with compact support is denoted by $C_0(\Omega)$ or $\mathcal{K}(\Omega)$, where the **support** (or **carrier**) of a function f is the \dagger closure of the set $\{x | f(x) \neq 0\}$ in Ω and is usually denoted by $\text{supp } f$. If Ω is not compact, $C_0(\Omega)$ is not complete with respect to the supremum norm, but when Ω is σ -compact $C_0(\Omega)$ is complete with respect to the strongest locally convex topology with the property that for each compact set K in Ω the embed-

ding in $C_0(\Omega)$ of the linear subspace \mathcal{K}_K of all functions with support in K equipped with the supremum norm is continuous. Usually $\mathcal{K}(\Omega)$ denotes the space $C_0(\Omega)$ equipped with this topology.

(2) **The Lebesgue Spaces $L_p(\Omega)$** ($0 < p < \infty$). Let (Ω, μ) be a \dagger measure space. We denote by $L_p(\Omega)$ the totality of \dagger measurable functions $f(x)$ on Ω such that $|f(x)|^p$ is integrable. A function f defined on (Ω, μ) is called **square integrable** if $f \in L_2(\Omega, \mu)$. If Ω is the interval (a, b) equipped with the Lebesgue measure, it is sometimes denoted by $L_p(a, b)$. We define the norm $\|f\|_p$ by

$$\|f\| = \|f\|_p = \left(\int_{\Omega} |f(x)|^p d\mu(x) \right)^{1/p}.$$

If $1 \leq p < \infty$, then $L_p(\Omega)$ is a Banach space. The triangle inequality for this norm is precisely the \dagger Minkowski inequality. If $0 < p < 1$, the norm no longer satisfies the triangle inequality but does satisfy the quasinorm inequality $\|f + g\|_p \leq 2^{1/p-1}(\|f\|_p + \|g\|_p)$, and if $\sum \|f_n\|_p < \infty$, then $\sum f_n$ converges unconditionally in $L_p(\Omega)$. Hence $L_p(\Omega)$, $0 < p < 1$, is a \dagger quasi-Banach space. If $\lim_{n \rightarrow \infty} \|f_n - f\|_p = 0$, we say that the sequence $\{f_n\}$ **converges to f in the mean of order p** (or **in the mean of power p**), and write $\text{l.i.m.}_{n \rightarrow \infty} f_n = f$. If $\{f_n\}$ converges to f in the mean of order 2, we simply say that $\{f_n\}$ converges to f in the mean. (The notation l.i.m. means the **limit in the mean** and is used mostly when $p = 2$.) For any $f, g \in L_2(\Omega)$, $(f, g) = \int_{\Omega} f(x) \overline{g(x)} d\mu(x)$ is well defined, by the \dagger Schwarz inequality, and has the properties of the \dagger inner product. Hence, $L_2(\Omega)$ is a \dagger Hilbert space. If $1 < p < \infty$, then $L_p(\Omega)$ is a \dagger uniformly convex Banach space and is in particular \dagger reflexive. Deepest results on $L_p(\mathbb{R}^n)$, $1 < p < \infty$, are often derived from the **Littlewood-Paley theory** due to J. E. Littlewood and R. E. A. C. Paley, A. Zygmund [6], and E. M. Stein [7]. Its starting point is the inequality

$$A_p \|f\|_p \leq \|g(f)\|_p \leq A'_p \|f\|_p,$$

where $g(f)$ is the function

$$g(f)(x) = \left(\int_0^\infty |\text{grad}_x u(x, t)|^2 t dt \right)^{1/2}$$

obtained from the **Poisson integral**

$$u(x, t)$$

$$= \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \int_{\mathbb{R}^n} t(t^2 + |x - y|^2)^{-(n+1)/2} f(y) dy,$$

where $|x| = (x_1^2 + \dots + x_n^2)^{1/2}$.

The L_p spaces, $1 < p < \infty$, are generalized in the following way. Let $\Phi(s)$ be a convex and nondecreasing function on $[0, \infty)$ satisfying $\Phi(0) = 0$ and $\Phi(s)/s \rightarrow \infty$ as $s \rightarrow \infty$. Denote

by $L_\Phi(\Omega)$ ($L_\Phi^*(\Omega)$) the set of all functions $f(x)$ such that $\Phi(|f(x)|)$ is integrable ($\Phi(k|f(x)|)$ is integrable for some $k > 0$). $L_\Phi(\Omega) = L_\Phi^*(\Omega)$ if $\Phi(2s) \leq C\Phi(s)$. $L_\Phi^*(\Omega)$ is a Banach space, called the **Orlicz space**, under the norm

$$\|f\| = \inf \left\{ \lambda > 0 \mid \int \Phi(\lambda^{-1}|f(x)|) d\mu(x) \leq 1 \right\}.$$

$L_p(\Omega)$, $1 < p < \infty$, is the Orlicz space for $\Phi(s) = s^p$.

(3) The Function Space $M(\Omega) = L_\infty(\Omega)$. Let Ω and μ be as in (2). A measurable function $f(x)$ on Ω is said to be **essentially bounded** if there exists a positive number α such that $|f(x)| \leq \alpha$ almost everywhere on Ω . The infimum of such α is called the **essential supremum** of f , denoted by $\text{ess sup}_{x \in \Omega} |f(x)|$. The totality of essentially bounded measurable functions on Ω (denoted by $M(\Omega)$) is a Banach space with respect to the norm $\|f\| = \|f\|_\infty = \text{ess sup}_{x \in \Omega} |f(x)|$. If $\mu(\Omega) < \infty$, then $M(\Omega) \subset L_p(\Omega)$ for any $p > 0$, and $\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p$ for any $f \in M(\Omega)$. From this point of view, $M(\Omega)$ is also denoted by $L_\infty(\Omega)$ even when $\mu(\Omega) = \infty$. This is also the reason why the notation $\|\cdot\|_\infty$ is used for the norm in $M(\Omega)$.

(4) The Lorentz Spaces $L_{(p,q)}(\Omega)$ ($0 < p, q \leq \infty$). The Lebesgue spaces $L_p(\Omega)$, $0 < p \leq \infty$, are **rearrangement invariant**. Namely, define for a measurable function f on a measure space (Ω, μ) the **distribution function** $\mu_f(s) = \mu\{x \in \Omega \mid |f(x)| > s\}$, $s > 0$, and the **rearrangement** $f^*(t) = \inf\{s > 0 \mid \mu_f(s) \leq t\}$, $t > 0$. Then $f \in L_p(\Omega)$ if and only if $f^* \in L_p(0, \infty)$ and $\|f\|_p = \|f^*\|_p$. Another important class of rearrangement invariant spaces are the **Lorentz spaces** $L_{(p,q)}(\Omega)$, $0 < p, q \leq \infty$ (G. G. Lorentz, 1950; R. A. Hunt [8]), which is defined to be the quasi-Banach space of all measurable functions f on Ω such that

$$\|f\|_{(p,q)} = \|t^{1/p} f^*(t)\|_{L_q^*} < \infty,$$

where L_q^* is the L_q -space on $(0, \infty)$ relative to the measure dt/t . $L_{(p,p)}(\Omega) = L_p(\Omega)$ with equal norms. If $1 < p < \infty$ and $1 \leq q \leq \infty$, then $L_{(p,q)}(\Omega)$ is a Banach space under the equivalent norm $\|t^{1/p-1} \int_0^t f^*(s) ds\|_{L_q^*}$. Except for these cases, $L_{(p,q)}(\Omega)$ is not equivalent to a normed space in general [8]. If $q_0 < q_1$, then $L_{(p,q_0)}(\Omega) \subset L_{(p,q_1)}(\Omega)$ with continuous embedding. In case $\mu(\Omega) < \infty$, $L_{(p_0,q_0)}(\Omega) \subset L_{(p_1,q_1)}(\Omega)$ for $p_0 > p_1$ and any q_0, q_1 . The Lorentz spaces play an important role in interpolation and approximation theory (\rightarrow 224 Interpolation of Operators).

(5) The Function Space $S(\Omega)$. Let (Ω, μ) be a measure space with $\mu(\Omega) < \infty$. Denote by $S(\Omega)$

the totality of measurable functions on Ω that take finite value almost everywhere. Then $\|f\| = \int_\Omega (|f(x)|/(1+|f(x)|)) d\mu(x)$ for $f \in S(\Omega)$ has the properties of the \dagger pseudonorm and $S(\Omega)$ is a \dagger Fréchet space (in the sense of Banach) that is not locally convex in general. We have $\lim_{n \rightarrow \infty} \|f_n - f\| = 0$ if and only if

$$\lim_{n \rightarrow \infty} \mu(\{x \mid |f_n(x) - f(x)| > \varepsilon\}) = 0$$

for any positive number ε . Convergence of this type is called **convergence in measure** (or **asymptotic convergence**) and is the same notion as \dagger convergence in probability of a sequence of \dagger random variables (\rightarrow 342 Probability Theory). If $\{f_n\} \in L_p(\Omega)$ converges to $f \in L_p(\Omega)$ in the mean of order p , then $\{f_n\}$ converges to f asymptotically, but the converse is not true in general. If a sequence $\{f_n\} \in S(\Omega)$ converges to $f \in S(\Omega)$ almost everywhere, then $\{f_n\}$ converges to f asymptotically. Any sequence $\{f_n\}$ that converges to f asymptotically contains a subsequence $\{f_{n_k}\}$ that converges to f almost everywhere.

(6) The Sequence Spaces c, c_0, l_p ($0 < p < \infty$), $m = l_\infty$, and s . The totality c (resp. c_0) of sequences $x = \{\xi_n\}$ that converges (resp. converges to zero) as $n \rightarrow \infty$ forms a Banach space with respect to the norm $\|x\| = \sup |\xi_n|$. c_0 (resp. c) is the space $C_\infty(\Omega)$ (resp. $C(\Omega)$) when Ω is (resp. the one-point compactification of) the discrete locally compact space $\{1, 2, 3, \dots\}$. The sequence space l_p , $0 < p < \infty$ (resp. $m = l_\infty$), is defined to be the spaces $L_p(\Omega)$ (resp. $M(\Omega)$), where Ω is the space $\{1, 2, 3, \dots, n, \dots\}$, of which each point has unit mass, while s denotes the space $S(\Omega)$, where $\Omega = \{1, 2, \dots, n, \dots\}$, provided with the measure assigning mass $1/2^n$ to the point n . s is the set of all sequences equipped with the topology of pointwise convergence (s is also used to denote the space of rapidly decreasing sequences; \rightarrow Section (16)).

Assume that the space $L_2(\Omega)$ mentioned in (2) is \dagger separable and that $\{\varphi_n\}$ is a \dagger complete orthonormal set in $L_2(\Omega)$. Then putting

$$\xi_n = \int_\Omega f(x) \overline{\varphi_n(x)} d\mu(x)$$

(\dagger Fourier coefficients) for any $f \in L_2(\Omega)$, we have $\{\xi_n\} \in l_2$ and $\sum_{n=1}^\infty |\xi_n|^2 = \|f\|^2$. Conversely, for any $\{\xi_n\} \in l_2$ there exists an $f = \sum_{n=1}^\infty \xi_n \varphi_n \in L_2(\Omega)$ whose Fourier coefficients are the given ξ_n (**Riesz-Fischer theorem**). By means of this correspondence, separable spaces $L_2(\Omega)$ and l_2 are mutually isomorphic as Hilbert spaces. Sometimes we denote by $l_p(\Omega)$ the function space $L_p(\Omega)$, where Ω is an arbitrary set endowed with the measure assigning mass 1 to each point.

(7) **The John-Nirenberg Space BMO .** A locally integrable function $f(x)$ on \mathbf{R}^n is said to be of **bounded mean oscillation** if

$$\|f\|_{BMO} = \sup |S|^{-1} \int_S |f(x) - f_S| dx < \infty, \quad (1)$$

where the supremum is taken over all (solid) spheres S in \mathbf{R}^n , f_S is the mean $|S|^{-1} \int_S f(x) dx$, and $|S|$ denotes the measure of S (F. John and L. Nirenberg, 1961). The set $BMO(\mathbf{R}^n)$ of all functions on \mathbf{R}^n of bounded mean oscillation forms a Banach space under the norm mentioned above if two functions f and g are identified whenever $f - g$ is equal to a constant almost everywhere. Condition (1) is equivalent to

$$\sup \left(|S|^{-1} \int_S |f(x) - f_S|^p dx \right)^{1/p} < \infty$$

for any $1 \leq p < \infty$. There are constants B and $K > 0$ such that

$$\begin{aligned} |\{x \in S \mid |f(x) - f_S| > \lambda\}| \\ \leq B|S| \exp(-K\lambda/\|f\|_{BMO}) \end{aligned}$$

for any sphere S and $\lambda > 0$. BMO is a slightly larger space than L_∞ (e.g. $\log|x| \in BMO$) and has better properties. For example, the \dagger Calderón-Zygmund operators are bounded in $BMO(\mathbf{R}^n)$. We have

$$BMO(\mathbf{R}^n) = L_\infty(\mathbf{R}^n) + \sum_{j=1}^n R_j L_\infty(\mathbf{R}^n),$$

where R_j are the \dagger Riesz transforms [9]. This is called the **Fefferman-Stein decomposition**. The Riesz transforms can be replaced by more general families of singular integral operators (A. Uchiyama, *Acta Math.* (1982)).

(8) **The Hardy Spaces H_p** ($0 < p < \infty$). The classical theory of Hardy classes (\rightarrow 159 Fourier Series G) has been reconstructed by the real-analysis method and extended to higher-dimensional cases by E. M. Stein, G. Weiss, C. Fefferman, and others. According to their terminology the elements of the Hardy space H_p are (the complex linear combinations of) the real parts of the boundary values of holomorphic functions of the Hardy class.

Let $f \in \mathcal{S}'(\mathbf{R}^n)$ be a \dagger tempered distribution. For a $\varphi \in \mathcal{S}(\mathbf{R}^n)$ define the radial maximal function $M_\varphi^+ f$ and the nontangential maximal function $M_\varphi^* f$ relative to φ by

$$M_\varphi^+ f(x) = \sup_{t>0} |\varphi_t * f(x)|,$$

$$M_\varphi^* f(x) = \sup_{|x-y| \leq t} |\varphi_t * f(y)|,$$

where $\varphi_t(x) = t^{-n} \varphi(x/t)$ and $*$ denotes convolution. Then the **Hardy space $H_p(\mathbf{R}^n)$** , $0 < p < \infty$, is defined to be the space of all tempered distributions f which satisfy the following

equivalent conditions: (i) $M_\varphi^+ f \in L_p$ for a φ with $\int \varphi(x) dx \neq 0$; (ii) $M_\varphi^* f \in L_p$ for a φ with $\int \varphi dx \neq 0$; (iii) $M_\varphi^+ f \in L_p$ for any φ ; (iv) $M_\varphi^* f \in L_p$ for any φ . If a distribution f satisfies (one of) these conditions, then its Poisson integral $u(x, t) = \varphi_t * f(x)$ is a function, where $\varphi(x) = \text{const}(1 + |x|^2)^{-(n+1)/2}$, and its **radial maximal function** $u^+(x) = \sup_{t>0} |u(x, t)|$ and **nontangential maximal function** $u^*(x) = \sup_{|x-y| < t} |u(y, t)|$ both belong to L_p . Conversely if $u(x, t)$ is a harmonic function on the upper half-space $t > 0$ and if either u^+ or u^* belongs to L_p , then its boundary value $f = u(\cdot, 0)$ exists in the sense of tempered distribution and f satisfies the above conditions. Define the norm of an $f \in H_p(\mathbf{R}^n)$ by $\|u^*\|_p$. Then $H_p(\mathbf{R}^n)$ becomes a quasi-Banach space. If $1 < p < \infty$, then $H_p(\mathbf{R}^n) = L_p(\mathbf{R}^n)$ with equivalent norms. $H_1(\mathbf{R}^n)$ is a Banach space strictly smaller than $L_1(\mathbf{R}^n)$. An $f \in L_1(\mathbf{R}^n)$ belongs to $H_1(\mathbf{R}^n)$ if and only if all the Riesz transforms $R_j f$ are in $L_1(\mathbf{R}^n)$, and $\|f\|_{H_1}$ is equivalent to $\|f\|_1 + \sum \|R_j f\|_1$. Similar characterizations of $H_p(\mathbf{R}^n)$ are also known for $p > 0$ (Fefferman and Stein [9]).

Let $\Gamma_j, j = 1, \dots, m$, be \dagger proper convex open cones in \mathbf{R}^n such that the \dagger polars Γ_j° cover \mathbf{R}^n . Then an $f \in \mathcal{S}'(\mathbf{R}^n)$ belongs to $H_p(\mathbf{R}^n)$ if and only if there are holomorphic functions $F_j(x + iy)$ on $\mathbf{R}^n + i\Gamma_j$ such that $\sup \{\|F_j(\cdot + iy)\|_p \mid y \in \Gamma_j\} < \infty$ and $f = \sum F_j(\cdot + i\Gamma_j 0)$ (D. L. Burkholder, R. F. Gundy, and M. L. Silverstein for $n = 1$ and L. Carleson for $n > 1$). Let $0 < p \leq 1$. A measurable function a on \mathbf{R}^n is said to be a **p -atom** if there is a sphere S such that $\text{supp } a \subset S$ and $\|a\|_\infty \leq |S|^{-1/p}$ and if $\int a(x) x^\alpha dx = 0$ for all multi-indices α with $|\alpha| \leq n(p^{-1} - 1)$. Here a **multi-index** α is an n -tuple $(\alpha_1, \dots, \alpha_n)$ of nonnegative integers, $|\alpha| = \alpha_1 + \dots + \alpha_n$ and $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$. A distribution f belongs to $H_p(\mathbf{R}^n)$, $0 < p \leq 1$, if and only if there are a sequence of p -atoms a_j and a sequence of numbers $\lambda_j \geq 0$ in l_p such that $f = \sum \lambda_j a_j$ in the sense of distributions, and the norm $\|f\|_{H_p}$ is equivalent to the infimum of $\|\lambda_j\|_{l_p}$ (R. R. Coifman for $n = 1$ and R. H. Latter for $n > 1$). The theory of H_p and BMO has been generalized to more general situations (\rightarrow Coifman and Weiss, *Bull. Amer. Math. Soc.* (1977)).

From now on we assume that Ω is a domain in the n -dimensional Euclidean space \mathbf{R}^n (or more generally a differentiable manifold). D^α stands for $D_1^{\alpha_1} \dots D_n^{\alpha_n}$, where $D_j = \partial/\partial x_j$.

(9) **The Function Spaces $C^l(\Omega)$ and $C_0^l(\Omega)$** ($l = 0, 1, 2, \dots, \infty$). The totality of l -times continuously differentiable functions in Ω (namely, differentiable functions of \dagger class C^l in Ω) is denoted by $C^l(\Omega)$. We say that a sequence $\{f_v\}$ of functions in $C^l(\Omega)$ converges to 0 in $C^l(\Omega)$ if $|D^\alpha f_v(x)|$ converges to 0 uniformly on

every compact subset of Ω for every α satisfying $0 \leq |\alpha| \leq l$ ($0 \leq |\alpha| < \infty$ if $l = \infty$). $C^l(\Omega)$ is a Fréchet space. The totality of functions in $C^l(\Omega)$ whose supports are compact subsets of Ω is denoted by $C_0^l(\Omega)$. We say that a sequence $\{f_v\}$ of functions in $C_0^l(\Omega)$ converges to 0 in $C_0^l(\Omega)$ if $\text{supp } f_v$ ($v = 1, 2, \dots$) is contained in a compact subset of Ω independent of v and $\{f_v\}$ converges to 0 in $C^l(\Omega)$. $C_0^l(\Omega)$ is an ${}^t(\text{LF})$ -space.

When Ω is a locally closed set in \mathbf{R}^n (or a differentiable manifold with boundary), we denote by $C^l(\Omega)$ the totality of functions $f(x)$ on Ω together with their continuous formal derivatives $D^\alpha f(x)$, $|\alpha| \leq l$ ($|\alpha| < \infty$ if $l = \infty$) on Ω such that for every α and m with $|\alpha| \leq m \leq l$ ($m < \infty$ if $l = \infty$),

$$D^\alpha f(x) - \sum_{|\beta| \leq m - |\alpha|} D^{\alpha+\beta} f(y) (x-y)^\beta / \beta! \\ = o(|x-y|^{m-|\alpha|})$$

locally uniformly in Ω as $|x-y|$ tends to 0. Convergence in $C^l(\Omega)$ is defined in the same way as above.

If Ω is a closed set in \mathbf{R}^n with a finite number of connected components in each of which two points x and y are connected by an arc of length $\leq C|x-y|$ with a constant C independent of x and y , then every f in $C^l(\Omega)$ can be extended to an \tilde{f} in $C^l(\mathbf{R}^n)$ (**Whitney's extension theorem**).

(10) The Lipschitz Spaces Λ^s . Let $s > 0$, and let k be the least integer greater than s . The **Lipschitz** (or **Hölder**) **space $\Lambda^s(\mathbf{R}^n)$** is the totality of functions $f(x)$ on \mathbf{R}^n which satisfy

$$\|f\|_{\Lambda^s} = \sup_{x,y} \left| \sum_{j=0}^k \binom{k}{j} (-1)^j f(x+jy) \right| / |y|^s < \infty.$$

When $s < 1$, this is exactly the t Lipschitz (or Hölder) condition of order s . But when $s = 1$, it is strictly weaker than the t Lipschitz condition. A function $f \in \Lambda^1$ is said to be **smooth in the sense of A. Zygmund** [6]. Suppose $0 < h < s$ is an integer. Then a function f belongs to Λ^s if and only if it is h times continuously differentiable and all the derivatives $D^\alpha f$ of order h are in Λ^{s-h} . $\Lambda^s(\mathbf{R}^n)$ is a Banach space of functions modulo the polynomials of degree $\leq k-1$.

Suppose that $1 \leq q < \infty$ and f is a measurable function on \mathbf{R}^n such that

$$\sup_S \inf_P \left(|S|^{-1-(sq/n)} \int_S |f(x) - P(x)|^q dx \right)^{1/q} < \infty,$$

where the supremum is taken over all spheres in \mathbf{R}^n and the infimum over all polynomials of degree $\leq s$. Then $f(x)$ is equal to a function $\tilde{f}(x)$ in $\Lambda^s(\mathbf{R}^n)$ almost everywhere and the supremum is equivalent to the norm $\|\tilde{f}\|_{\Lambda^s}$. Conversely every $f \in \Lambda^s(\mathbf{R}^n)$ satisfies the above inequality (S. Campanato).

(11) The Sobolev Spaces $W_p^l(\Omega)$, $H^l(\Omega)$, and $H_0^l(\Omega)$ ($l = 0, 1, 2, \dots, 1 \leq p \leq \infty$ or $-\infty < l < \infty, 1 < p < \infty$). Let $l \geq 0$ be an integer and $1 \leq p \leq \infty$. The Sobolev space $W_p^l(\Omega)$ is the totality of functions $f(x)$ such that for all α satisfying $|\alpha| \leq l$, the derivatives $D^\alpha f(x)$ in the sense of distribution (\rightarrow 125 Distributions and Hyperfunctions) belong to $L_p(\Omega)$ with respect to Lebesgue measure in Ω . $W_p^l(\Omega)$ is a Banach space with the norm

$$\|f\| = \|f\|_{W_p^l} = \left(\sum_{0 \leq |\alpha| \leq l} \int_\Omega |D^\alpha f(x)|^p dx \right)^{1/p}.$$

Clearly $W_p^0(\Omega) = L_p(\Omega)$. $W_2^l(\Omega)$ is a Hilbert space with respect to the inner product

$$(f, g) = \int_\Omega \sum_{0 \leq |\alpha| \leq l} D^\alpha f(x) \cdot \overline{D^\alpha g(x)} dx.$$

Sometimes $W_2^l(\Omega)$ is denoted by $H^l(\Omega)$. Its closed linear subspace obtained as the completion of $C_0^\infty(\Omega)$ is denoted by $H_0^l(\Omega)$. We have $H_0^0(\Omega) = W_2^0(\Omega) = L_2(\Omega)$. However, if $l \geq 1$, we have $H_0^l(\Omega) \subset W_2^l(\Omega)$, and identity does not hold unless $\Omega = \mathbf{R}^n$.

The definition of Sobolev spaces has been extended to those with fractional and also negative order $-\infty < s < \infty$ in many different ways. When $1 < p < \infty$, it is natural to define $W_p^s(\mathbf{R}^n)$ to be the space of all tempered distributions f on \mathbf{R}^n such that $(1 - \Delta)^{s/2} f = \mathcal{F}^{-1}((1 + |\xi|^2)^{s/2} \mathcal{F}f) \in L_p(\mathbf{R}^n)$. If $s > 0$, then $W_p^s(\mathbf{R}^n)$ thus defined coincides with the space of all $f \in L_p(\mathbf{R}^n)$ whose Poisson integral $u(x, t)$ satisfies

$$\left\| \left(\int_0^\infty t^{4k-2s-1} |\Delta^k u(\cdot, t)|^2 dt \right)^{1/2} \right\|_p < \infty$$

for some (and any) integer $k > s/2$.

The Poisson integral can be replaced by other regularizations, and thus the definition of Sobolev spaces of fractional orders is extended to arbitrary open set Ω with the cone condition (T. Muramatsu [13]). Here Ω is said to satisfy the cone condition if there are a bounded and uniformly continuous mapping $\Psi: \mathbf{R}^n \rightarrow \mathbf{R}^n$ and an $\varepsilon > 0$ such that for any $x \in \Omega$ the convex hull of the ε -ball with center at $x + \Psi(x)$ and $\{x\}$ is included in Ω .

(12) The Besov Spaces $B_{p,q}^s$ ($-\infty < s < \infty, 1 \leq p, q \leq \infty$). The effort to make the Sobolev embedding theorem [10] more precise led S. M. Nikol'skii and O. V. Besov [11] to the other classes of "Sobolev spaces" of fractional order. Let $s > 0$ and $1 \leq p, q \leq \infty$. The **Besov space $B_{p,q}^s(\mathbf{R}^n)$** is the totality of functions $f \in L_p(\mathbf{R}^n)$ such that

$$\|f\|_{B_{p,q}^s} = \left(\int \left\| \sum_{j=0}^k \binom{k}{j} (-1)^j f(\cdot + jy) \right\|_p^q \frac{dy}{|y|^{qs+n}} \right)^{1/q} < \infty$$

for some (and any) integer $k > s$. In terms of the Poisson integral an $f \in L_p(\mathbf{R}^n)$ belongs to $B_{p,q}^s(\mathbf{R}^n)$ if and only if

$$\left(\int_0^\infty t^{(2k-s)q-1} \left\| \Delta^k u(\cdot, t) \right\|_p^q dt \right)^{1/p} < \infty$$

for some (and any) integer $k > s/2$. The first definition is obviously extended to an arbitrary domain Ω in \mathbf{R}^n . If Ω satisfies the cone condition, then the functions in $B_{p,q}^s(\Omega)$ are characterized similarly as above by using a suitable regularization $u(x, t)$ of $f(x)$ (Muramatu [13]). Let Ω be a domain with the cone condition. When $0 < h < s$ is an integer, an f belongs to $B_{p,q}^s(\Omega)$ if and only if $f \in W_p^h(\Omega)$ and all the derivatives $D^\alpha f$ of order h are in $B_{p,q}^{s-h}(\Omega)$. $B_{p,q}^s(\Omega)$ is a Banach space with the norm $\|f\|_p + |f|_{B_{p,q}^s}$. If $p=2$, then $B_{2,2}^s(\Omega)$ coincides with the Sobolev space $W_2^s(\Omega)$. But in the other cases, $B_{p,q}^s(\Omega)$ is different from $W_p^s(\Omega)$ for any q . However, $B_{p,p}^s(\Omega)$ or $B_{p,\infty}^s(\Omega)$ is often called a Sobolev space of order s and denoted by $W_p^s(\Omega)$. Clearly we have $B_{\infty,\infty}^s = L_\infty \cap \Lambda^s$.

The **Besov space** $B_{p,q}^s(\Omega)$ of order $s \leq 0$ is defined to be the totality of distributions f which can be represented as $f = \sum_{|\alpha| \leq k} D^\alpha f_\alpha$ with $f_\alpha \in B_{p,q}^{s+k}(\Omega)$ for some (and any) integer k such that $s+k > 0$. The norm is defined to be $\inf \sum_\alpha \|f_\alpha\|_{B_{p,q}^{s+k}(\Omega)}$. In terms of the Poisson integrals or other regularizations the same characterization holds as above.

If Ω is a domain in \mathbf{R}^n with the cone condition, then the restriction $B_{p,q}^s(\mathbf{R}^n) \rightarrow B_{p,q}^s(\Omega)$ (resp. $W_p^s(\mathbf{R}^n) \rightarrow W_p^s(\Omega)$ for $1 < p < \infty$) is a bounded linear surjection with a bounded right inverse [13].

From now on we denote by $B_{p,q}^s$, etc., $B_{p,q}^s(\Omega)$, etc. for a domain $\Omega \subset \mathbf{R}^n$ with the cone condition. If $q < r$, then $B_{p,q}^s \subset B_{p,r}^s$. If $1 \leq p \leq 2$, then $B_{p,p}^s \subset W_p^s \subset B_{p,2}^s$. If $2 \leq p \leq \infty$, then $B_{p,2}^s \subset W_p^s \subset B_{p,p}^s$. If $s > t$, then $B_{p,\infty}^s \subset B_{p,\infty}^t$.

Sobolev-Besov embedding theorems: (i) Let $p < p'$ and $s - n/p = s' - n/p'$. Then $B_{p,q}^s \subset B_{p',q}^{s'}$, $B_{p,p'}^s \subset W_{p'}^{s'}$, $W_p^s \subset B_{p',p}^{s'}$ and, if $p' < \infty$, then $W_{p'}^{s'} \subset W_p^{s'}$. (ii) Let $0 < s < n/p$ and $n/p' = n/p - s$. Then $B_{p,q}^s \subset L_{(p',q)}$ and $W_p^s \subset L_{(p',p)} (\subset L_{p'})$. (iii) Let $s = n/p$. Then $B_{p,1}^s \subset BC$ for any p and $B_{p,\infty}^s \subset BMO$ for $p < \infty$. (iv) Let $0 \leq n' < n$ and $0 < s' = s - (n - n')/p$. Then there is a bounded **trace operator** $Tr: B_{p,q}^s(\mathbf{R}^n) \rightarrow B_{p,q}^{s'}(\mathbf{R}^{n'})$ that extends the restriction mapping $\mathcal{S}(\mathbf{R}^n) \rightarrow \mathcal{S}(\mathbf{R}^{n'})$. Tr is surjective and has a bounded linear right inverse [11–13].

(13) The Function Spaces \mathcal{D} , \mathcal{E} , \mathcal{D}_{L_p} , \mathcal{B} , and \mathcal{S} . The spaces of infinitely differentiable functions $C_0^\infty(\Omega)$ and $C^\infty(\Omega)$ are also denoted by $\mathcal{D}(\Omega)$ and $\mathcal{E}(\Omega)$, respectively. The totality of functions $f(x)$ in $C^\infty(\Omega)$ such that, for all α , $D^\alpha f(x)$ belongs to $L_p(\Omega)$ with respect to † Lebesgue measure is denoted by $\mathcal{D}_{L_p}(\Omega)$. The neighbor-

hood $V_{l,\varepsilon}$ of 0 in $\mathcal{D}_{L_p}(\Omega)$ is defined to be the totality of functions $f(x)$ such that $\|D^\alpha f\|_p < \varepsilon$ for any α satisfying $|\alpha| \leq l$. In particular, $\mathcal{D}_{L_\infty}(\Omega)$ is also denoted by $\mathcal{B}(\Omega)$. ($\mathcal{B}(\Omega)$ is also used to denote the space of hyperfunctions.)

A function $f(x)$ is called a **rapidly decreasing C^∞ -function** if it belongs to $C^\infty(\mathbf{R}^n)$ and satisfies

$$\lim_{|x| \rightarrow \infty} |x|^k |D^\alpha f(x)| = 0$$

for any α and any integer $k > 0$. The totality of rapidly decreasing C^∞ -functions is denoted by \mathcal{S} . The neighborhood $V_{l,k,\varepsilon}$ of 0 in \mathcal{S} is defined to be the totality of functions $f(x)$ such that $(1 + |x|^2)^k |D^\alpha f(x)| < \varepsilon$ for any α satisfying $|\alpha| \leq l$.

The spaces in this section are † nuclear except for \mathcal{D}_L and \mathcal{B} , and are employed in the theory of distributions [14] (\rightarrow 125 Distributions and Hyperfunctions). When $\Omega = \mathbf{R}^n$, we usually omit (\mathbf{R}^n) ; for example, $\mathcal{D}(\mathbf{R}^n)$ and $\mathcal{D}_{L_p}(\mathbf{R}^n)$ are denoted by \mathcal{D} and \mathcal{D}_{L_p} , respectively.

(14) The Function Spaces $\mathcal{D}_{\{M_p\}}$, $\mathcal{D}_{(M_p)}$, $\mathcal{E}_{\{M_p\}}$, $\mathcal{E}_{(M_p)}$, and $\mathcal{A} = C^\omega$. Let $\{M_p\}$ be a sequence of positive numbers satisfying the logarithmic convexity $M_p^2 \leq M_{p-1} M_{p+1}$. $\mathcal{E}_{\{M_p\}}(\Omega)$ (resp. $\mathcal{E}_{(M_p)}(\Omega)$) denotes the totality of C^∞ -functions f on Ω such that for any compact set K in Ω there are constants k and C (resp. for any $k > 0$ there is a constant C) satisfying

$$|D^\alpha f(x)| \leq C k^{|\alpha|} M_{|\alpha|}, \quad x \in K, |\alpha| \geq 0.$$

$\mathcal{E}_{\{p\}}(\Omega)$ is the totality of † real analytic functions on Ω and is denoted by $\mathcal{A}(\Omega)$ or $C^\omega(\Omega)$. If $\{M_p\}$ satisfies the **Denjoy-Carleman condition** $\sum M_p/M_{p+1} < \infty$, then $\mathcal{D}_{\{M_p\}}(\Omega) = \mathcal{D}(\Omega) \cap \mathcal{E}_{\{M_p\}}(\Omega)$ and $\mathcal{D}_{(M_p)}(\Omega) = \mathcal{D}(\Omega) \cap \mathcal{E}_{(M_p)}(\Omega)$ are dense in $\mathcal{D}(\Omega)$. Conversely, if $\mathcal{D}_{\{M_p\}}(\Omega)$ is different from $\{0\}$, then $\{M_p\}$ satisfies the Denjoy-Carleman condition. In this case an $f \in \mathcal{E}_{\{M_p\}}(\Omega)$ (resp. $\mathcal{E}_{(M_p)}(\Omega)$) is sometimes called an **ultradifferentiable function** of class $\{M_p\}$ (resp. (M_p)). The most important is the case where $M_p = p!^s$ for an $s > 1$. Then an $f \in \mathcal{E}_{\{M_p\}}(\Omega)$ (resp. $\mathcal{E}_{(M_p)}(\Omega)$) is called a **function of Gevrey class $\{s\}$** (resp. (s)). The topological properties of $\mathcal{A}(\Omega)$ (resp. $\mathcal{E}_{\{M_p\}}(\Omega)$, etc.) have been discussed by A. Martineau (resp. H. Komatsu).

For function spaces of S type \rightarrow 125 Distributions and Hyperfunctions.

(15) The Function Spaces $\mathcal{O}(\Omega)$, $\mathcal{O}_p(\Omega) = A_p(\Omega)$, and $A(\Omega)$. Let Ω be an open set in \mathbf{C}^n . The totality of holomorphic functions on Ω is denoted by $\mathcal{O}(\Omega)$. $\mathcal{O}(\Omega)$ is a † nuclear Fréchet space with the topology of uniform convergence on compact sets. It is a closed linear subspace of $C(\Omega)$ and also of $C^\infty(\Omega)$.

For any $p \geq 1$, the totality of functions f

holomorphic in Ω and satisfying $\int_{\Omega} |f(z)|^p dx dy < \infty$ ($z = x + iy$; $dx dy$ is Lebesgue measure), denoted by $\mathcal{O}_p(\Omega)$ or $A_p(\Omega)$, is a Banach space with respect to the norm $\|f\|_p = (\int_{\Omega} |f(z)|^p dx dy)^{1/p}$. In particular, it is a Hilbert space when $p = 2$.

The totality of functions bounded and continuous on the closure of Ω and holomorphic in Ω (denoted by $A(\Omega)$) is a Banach space with respect to the norm $\|f\| = \sup_{z \in \Omega} |f(z)|$.

(16) The Köthe Spaces $\bigcap \lambda(x^{(k)})$ and $\sum \lambda^{\times}(x^{(k)})$. Let $\{\alpha^{(k)}\}$ be an increasing sequence of sequences $\alpha^{(k)} = (\alpha_1^{(k)}, \alpha_2^{(k)}, \dots)$ of positive numbers. The **echelon space** $\bigcap \lambda^{(\alpha^{(k)})}$ of G. Köthe [15] is the totality of sequences $x = (\xi_1, \xi_2, \dots)$ such that $p^{(k)}(x) = \sum \alpha_i^{(k)} |\xi_i| < \infty$ for any k . It is a Fréchet space with the topology determined by the seminorms $p^{(k)}$. The **co-echelon space** $\sum \lambda^{\times}(x^{(k)})$ is the totality of sequences $y = (\eta_1, \eta_2, \dots)$ such that $|\eta_i| \leq C \alpha_i^{(k)}$ for some C and k . The space s of **rapidly decreasing sequences** and the space s' of **slowly increasing sequences** are the echelon space and the co-echelon space for the sequence $\alpha_i^{(k)} = i^k$. If $\alpha_i^{(k)} = k^i$, then we obtain the space of power series with infinite radius of convergence and the space of convergent power series. More generally, let $\alpha = (\alpha_i)$ be a sequence of positive numbers. The echelon spaces for $\alpha_i^{(k)} = \exp(k\alpha_i)$ and $\exp(-k^{-1}\alpha_i)$ are called the **infinite type power series space** and the **finite type power series space** and are denoted by $\Lambda_{\infty}(\alpha)$ and $\Lambda_1(\alpha)$, respectively. Echelon spaces and co-echelon spaces have been employed to construct examples and counterexamples in the theory of †locally convex spaces by Köthe [15], Grothendieck, Y. and T. Kōmura, E. Dubinsky, D. Vogt, and others.

C. Dual Spaces

When Ω is a compact Hausdorff space, any bounded linear functional Φ on $C(\Omega)$ is expressed by the †Stieltjes integral

$$\Phi(f) = \int_{\Omega} f(x) d\varphi(x), \quad f \in C(\Omega), \tag{2}$$

with a †Radon measure φ , i.e., a (real- or complex-valued) †regular †countably additive set function defined on the Borel sets in Ω . Since φ is of bounded variation, the totality of such φ is denoted by $BV(\Omega)$. Conversely, any $\varphi \in BV(\Omega)$ gives a bounded linear functional on $C(\Omega)$ defined by (2), and $\|\Phi\|$ = the total variation of φ over Ω . Hence the dual space of $C(\Omega)$ is isomorphic to the Banach space $BV(\Omega)$ with the norm $\|\Phi\|$.

Let Ω be a locally compact Hausdorff space. Then the dual space of $C_{\infty}(\Omega)$ is again the

Banach space $BV(\Omega)$ of Radon measures of bounded variation. On the other hand, the dual space of $C_0(\Omega)$ is the space of all Radon measures on Ω . N. Bourbaki takes this fact as the definition of measure.

Any bounded linear functional Φ on $L_1(\Omega)$ is expressed as

$$\Phi(f) = \int_{\Omega} f(x) \varphi(x) d\mu(x), \quad f \in L_1(\Omega), \tag{3}$$

with a suitable $\varphi \in M(\Omega)$; and $\|\Phi\| = \|\varphi\|_{\infty}$.

Conversely, any $\varphi \in M(\Omega)$ defines a bounded linear functional on $L_1(\Omega)$ by means of (3). Accordingly, the dual space of $L_1(\Omega)$ is isomorphic to $M(\Omega)$.

The dual space of $L_p(\Omega)$ ($1 < p < \infty$) is isomorphic to $L_q(\Omega)$, where q is the real number defined by $(1/p) + (1/q) = 1$ (accordingly, $1 < q < \infty$) and is called the **conjugate exponent** of p . Any bounded linear functional on $L_p(\Omega)$ is expressible by the formula in (3) (where $f \in L_p(\Omega)$) with $\varphi \in L_q(\Omega)$, and $\|\Phi\| = \|\varphi\|_q$.

The dual space of $M(\Omega)$ is isomorphic to the normed linear space of all (real- or complex-valued) finitely additive set functions φ defined on all measurable sets in Ω , of bounded variation over Ω , and absolutely continuous with respect to the measure μ given in Ω (i.e., $\mu(N) = 0$ implies $\varphi(N) = 0$). If $1 < p < \infty$ and $1 \leq q < \infty$, then the dual space of $L_{(p,q)}(\Omega)$ is isomorphic to $L_{(p',q')}(\Omega)$, where p' and q' are conjugate exponents of p and q , respectively.

If Ω is **nonatomic** (i.e., Ω has no set of positive measure that cannot be decomposed into two subsets of positive measure), then no continuous linear functionals exist other than zero on $S(\Omega)$ and on $L_p(\Omega)$ and $L_{(p,q)}(\Omega)$ for $0 < p < 1$.

The sequence spaces c_0 , l_p ($1 \leq p < \infty$), m , and s (the space defined in Section B (6)) are special cases of $C_{\infty}(\Omega)$, $L_p(\Omega)$, $M(\Omega)$, and $S(\Omega)$, respectively. Hence their dual spaces can be described explicitly. For example, the dual space of c_0 (resp. l_1) is l_1 (resp. m), and if $1 < p < \infty$, the dual space of l_p is l_q (where $(1/p) + (1/q) = 1$). The coupling of $x = (\xi_n)$ and $y = (\eta_n)$ is given by $\sum \xi_n \eta_n$. The dual space of s is the totality of sequences (η_n) such that $\eta_n = 0$ except for a finite number of n .

Let $VMO(\mathbb{R}^n)$ denote the closure of $C_0(\mathbb{R}^n)$ in $BMO(\mathbb{R}^n)$. Then the dual space of $VMO(\mathbb{R}^n)$ is identified with $H_1(\mathbb{R}^n)$. This can be considered to be a generalization of the †F. and M. Riesz theorem. On the other hand, the dual space of $H_1(\mathbb{R}^n)$ is $BMO(\mathbb{R}^n)$ [9]. Hardy spaces $H_p(\mathbb{R}^n)$, $0 < p < 1$, are not locally convex but have sufficiently many bounded linear functionals, and their dual spaces are identified with Lipschitz spaces $\Lambda^s(\mathbb{R}^n)$, where $s = n(p^{-1} - 1)$.

Let $1 < p, q < \infty$. Then the dual space of $W_p^s(\mathbb{R}^n)$ is isomorphic to $W_p^{-s}(\mathbb{R}^n)$ and that of

$B_{p,q}^s(\mathbf{R}^n)$ to $B_{p',q'}^{-s}(\mathbf{R}^n)$, where p' and q' are conjugate exponents of p and q , respectively. The dual space \mathcal{D}' of \mathcal{D} is defined to be the space of \dagger distributions; the dual spaces of \mathcal{E} , \mathcal{D}_{L_p} , and \mathcal{S} are (algebraically) linear subspaces of \mathcal{D}' . Similarly, the dual spaces of $\mathcal{D}_{\{M_p\}}$ and $\mathcal{D}_{(M_p)}$ are called the spaces of \dagger ultradistributions of classes $\{M_p\}$ and (M_p) , respectively. The dual space of \mathcal{A} is identified with the space of \dagger hyperfunctions with compact support (\rightarrow 125 Distributions and Hyperfunctions). A continuous linear functional on $\mathcal{O}(\Omega)$ is called an **analytic functional**. For each analytic functional Φ there is a compact set $L \subset \Omega$ such that $|\Phi(f)| \leq C \sup_{z \in L} |f(z)|$. A compact set K is called a **porter** of Φ if every compact neighborhood L of K satisfies this condition. Porters are similar to supports of generalized functions, but an analytic functional does not necessarily have a smallest porter.

The dual space of an echelon space is the corresponding co-echelon space.

References

- [1] S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932 (Chelsea, 1963).
- [2] N. Dunford and J. T. Schwartz, *Linear operators*, Interscience, I, 1958; II, 1963; III, 1971.
- [3] K. Yosida, *Functional analysis*, Springer, 1965.
- [4] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces I, II*, *Erg. Math.*, 92 (1977), 97 (1979).
- [5] A. Grothendieck, *Critères de compacité dans les espaces fonctionnels généraux*, *Amer. J. Math.*, 74 (1952), 168–186.
- [6] A. Zygmund, *Trigonometric series*, Cambridge Univ. Press, second edition, 1959.
- [7] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, 1971.
- [8] R. A. Hunt, *On $L(p, q)$ spaces*, *Enseignement Math.*, (2) 12 (1966), 249–276.
- [9] C. Fefferman and E. M. Stein, *H^p spaces of several variables*, *Acta Math.*, 129 (1972), 137–193.
- [10] S. L. Sobolev, *Applications of functional analysis in mathematical physics*, *Amer. Math. Soc.*, 1963. (Original in Russian, 1950.)
- [11] O. V. Besov, *Investigations of a family of function spaces in connection with theorems of imbedding and extension*, *Amer. Math. Soc. Transl., Ser. 2*, 40 (1964), 85–126. (Original in Russian, 1961.)
- [12] H. Triebel, *Interpolation theory, function spaces, differential operators*, North-Holland, 1978.

- [13] T. Muramatu, *On Besov spaces and Sobolev spaces of generalized functions defined on a general region*, *Publ. Res. Inst. Math. Sci.*, 9 (1974), 325–396.
- [14] L. Schwartz, *Théorie des distributions*, Hermann, revised edition, 1966.
- [15] G. Köthe, *Topological vector spaces I*, Springer, 1969. (Original in German, 1966.)

169 (XI.19) Function-Theoretic Null Sets

A. General Remarks

By a **function-theoretic null set** we mean an exceptional set that appears in theorem as the one asserting that a certain property holds with a “small exception.” We give below some of the more important examples of exceptional sets. For simplicity, we limit ourselves to the n -dimensional Euclidean space \mathbf{R}^n ($n \geq 2$).

B. Sets of Harmonic Measure Zero

Denote by χ_E the characteristic function of a set E on the boundary ∂D of a bounded domain D in \mathbf{R}^n . We call the \dagger hypofunction H_{χ_E} and \dagger hyperfunction \bar{H}_{χ_E} (\rightarrow 120 Dirichlet Problem) the **inner** and **outer harmonic measures** of E (with respect to D), respectively. When they coincide, the function is called the **harmonic measure** of E . A necessary and sufficient condition for E to be of inner harmonic measure zero is that $u \leq 0$ hold in D whenever a \dagger subharmonic function u bounded above in D satisfies $\limsup u(P) \leq 0$ as P tends to any point of $\partial D - E$. This theorem implies the following uniqueness theorem: If h is bounded and harmonic in D , if E is a set of inner harmonic measure zero on ∂D , and if $h(P) \rightarrow 0$ as P tends to any point of $\partial D - E$, then $h \equiv 0$. A necessary and sufficient condition for E to be of outer harmonic measure zero is that there exist a positive \dagger superharmonic function v in D such that $v(P) \rightarrow \infty$ as P tends to any point of E . (Concerning the existence of a limit for a subharmonic or \dagger harmonic function at every boundary point except those on a set of harmonic measure zero, \rightarrow 193 Harmonic Functions and Subharmonic Functions.)

C. Sets of Capacity Zero

Although there are many kinds of capacity (\rightarrow 48 Capacity), here we consider only \dagger logarithmic capacity and α -capacity ($\alpha > 0$). Let K be a nonempty compact set in \mathbf{R}^n . Set

$W(K) = \inf_{\mu} \iint \overline{PQ}^{-\alpha} d\mu(P) d\mu(Q)$, where μ runs through the class of nonnegative \dagger Radon measures of total mass 1 supported by K , and write $C_{\alpha}(K) = (W(K))^{-1/\alpha}$. Define $C_{\alpha}(\emptyset) = 0$ for an empty set \emptyset . For a general set $E \subset \mathbf{R}^n$, define the inner capacity by $\sup_{K \subset E} C_{\alpha}(K)$ and the outer capacity by the infimum of the inner capacity of an open set containing E . When the inner and outer capacities coincide, the common value is called the α -**capacity** (or **capacity of order α**) of E and is denoted by $C_{\alpha}(E)$. We denote the logarithmic capacity of E by $C_0(E)$. In order that $C_0(K) = 0$ ($n = 2$) or the \dagger Newtonian capacity $C_{n-2}(K) = 0$ ($n \geq 3$) for a compact set K , it is necessary and sufficient that the harmonic measure of K with respect to $G - K$ vanish for any bounded domain G containing K . Then K is removable for any harmonic function that is bounded or has a finite \dagger Dirichlet integral in $G - K$. In general, K is said to be **removable** for a family F of functions if for any domain G containing K and $f \in F$ defined in $G - K$, there exists $g \in F$ defined in G such that $g = f$ in $G - K$. Let K be a compact set in \mathbf{R}^2 with $C_0(K) = 0$ and G be a domain containing K . Let f be a holomorphic function defined in $G - K$ for which every point of K is an \dagger essential singularity. Then the set of \dagger exceptional values for f at every point of K is of logarithmic capacity zero. If f is \dagger meromorphic in $|z| < 1$ and

$$\iint \frac{|f'|^2}{(1+|f|^2)^2} (1-|z|)^{\alpha} dx dy < \infty \quad (0 \leq \alpha < 1),$$

then f has a finite limit in any angular domain with a vertex at every point of $|z| = 1$ except for those belonging to a set of α -capacity zero (logarithmic capacity zero if $\alpha = 0$).

D. Hausdorff Measure

Let $\alpha > 0$ and a set E be given in \mathbf{R}^n . Denote by σ a covering of E by a countable number of balls with radii d_1, d_2, \dots , all of which are smaller than ε (> 0). As $\varepsilon \rightarrow 0$, $\inf_{\sigma} \sum_k d_k^{\alpha}$ increases. The limit is called the **Hausdorff measure** of E of dimension α and is denoted by $\Lambda_{\alpha}(E)$. In order for a compact set K to be removable for the family of harmonic functions defined in a bounded domain and satisfying the \dagger Hölder condition of order α , it is necessary and sufficient that $\Lambda_{n-2+\alpha}(K) = 0$. Next, suppose that K is a compact set in a plane and the complement G of K with respect to the plane is connected. Set $\|f\|_q = (\iint_G |f|^q dx dy)^{1/q}$ for f holomorphic in G and $q, 1 \leq q < \infty$, and $\|f\|_{\infty} = \sup_G |f|$. Denote by H^q the family of $f \neq 0$ with $\|f\|_q < \infty$. If p is defined by $1/p + 1/q = 1$, then $\Lambda_{2-p}(K) < \infty$ implies $H^q = \emptyset$ for $q, 2 < q < \infty$, and $\Lambda_1(K) = 0$

Function-Theoretic Null Sets

implies $H^{\infty} = \emptyset$. Moreover, $H^2 = \emptyset$ if and only if $C_0(K) = 0$, and $H^q = \emptyset$ implies $C_{2-p}(K) = 0$ for $q, 2 < q \leq \infty$ [1].

E. Null Sets Defined with Respect to Families of Functions

Conversely, L. V. Ahlfors and A. Beurling characterized the size of sets in a plane by means of families of functions [2]. Let D be a domain, and let f represent a holomorphic function in D . Fix a point z_0 in D . Set

$$\mathfrak{B} = \{f \mid |f| \leq 1\},$$

$$\mathfrak{D} = \left\{f \mid \iint_D |f'|^2 dx dy \leq \pi\right\},$$

$$\mathfrak{E} = \{f \mid \text{the area of } R^c \geq \pi\},$$

where R^c is the complement of the \dagger range R of $(f(z) - f(z_0))^{-1}$. Denote by $\mathfrak{B}, \mathfrak{D}, \mathfrak{E}$ the families consisting of constants and \dagger univalent functions in $\mathfrak{B}, \mathfrak{D}, \mathfrak{E}$, respectively. Use the notation \mathfrak{F} to represent any one of these six families, and define $M_{\mathfrak{F}} = M_{\mathfrak{F}}(z_0; D)$ by $\sup\{|f'(z_0)| \mid f \in \mathfrak{F}\}$. Then $M_{\mathfrak{B}} = M_{\mathfrak{E}} \geq M_{\mathfrak{D}} = M_{\mathfrak{D}\mathfrak{E}} \geq M_{\mathfrak{B}\mathfrak{E}} = M_{\mathfrak{D}\mathfrak{E}}$, and $M_{\mathfrak{F}}(z_0; D) = 0$ implies $M_{\mathfrak{F}}(z; D) = 0$ for any $z \in D$.

Denote by $N_{\mathfrak{F}}$ the class of compact sets K such that the complement K^c of K is connected and $M_{\mathfrak{F}}(z; K^c) = 0$. We call $K \in N_{\mathfrak{F}}$ a **null set of class $N_{\mathfrak{F}}$** . In order for K to be removable for \mathfrak{B} or \mathfrak{D} , it is necessary and sufficient that $K \in N_{\mathfrak{B}}$ or $\in N_{\mathfrak{D}}$, respectively. We in general have

$$\{K \mid C_0(K) = 0\} \subsetneq N_{\mathfrak{B}} \subsetneq N_{\mathfrak{D}} \subsetneq N_{\mathfrak{B}\mathfrak{E}}.$$

If $\Lambda_1(K) = 0$, then $K \in N_{\mathfrak{B}}$. There exists a set $K \in N_{\mathfrak{B}}$ with $\Lambda_1(K) > 0$ (A. G. Vitushkin, *Dokl. Akad. Nauk SSSR*, 127 (1959); J. Garnett, *Proc. Amer. Math. Soc.*, 21 (1970)). When K is a subset of an analytic arc A , $K \in N_{\mathfrak{B}}$ implies $\Lambda_1(K) = 0$, $N_{\mathfrak{D}}$ is equal to $N_{\mathfrak{B}\mathfrak{E}}$, and K belongs to $N_{\mathfrak{D}}$ if and only if $C_0(A) = C_0(A - K)$. If an \dagger analytic function has an essential singularity at every point of $K \in N_{\mathfrak{B}}$, then any compact subset of the set of exceptional values at every point of K belongs to $N_{\mathfrak{B}}$. A necessary and sufficient condition for $K \in N_{\mathfrak{D}}$ is either that the complement of any one-to-one \dagger conformal image of K^c be of plane measure zero or that any \dagger univalent analytic function in K^c be reduced to a \dagger linear fractional function. If the union of at most a countable number of $N_{\mathfrak{B}}$ or $N_{\mathfrak{D}}$ sets is compact, it belongs to the same class. But it not true for $N_{\mathfrak{B}\mathfrak{E}}$ sets [3].

F. Analytic Capacity

For a compact set K , let D_K be the unbounded connected component of K^c . The quantity

Function-Theoretic Null Sets

$M_{\mathbb{B}}(\infty, D_K)$ is called the **analytic capacity** of K and is denoted by $\alpha(K)$. If $\alpha(K) > 0$, there exists an extremal function $f_0(z) = \alpha(K)z^{-1} + \dots$, called the \dagger Ahlfors function, that maps D_K onto a covering surface of the unit disk. In general, $\alpha(K)$ is not greater than the logarithmic capacity $C_0(K)$. If K is a continuum, then $\alpha(K) = C_0(K)$, and $\alpha(K)$ is attained by and only by $f_0(z)$, which maps D_K onto $|w| < 1$ conformally and $z = \infty$ to $w = 0$. For a linear set K , $\alpha(K)$ is equal to a quarter of its length (C. Pommerenke, *Arch. Math.*, 11 (1960)). The concept of analytic capacity is basic for the theory of rational approximation on compact sets.

References

- [1] L. Carleson, Selected problems on exceptional sets, Van Nostrand, 1967.
- [2] L. V. Ahlfors and A. Beurling, Conformal invariants and function-theoretic null-sets, *Acta Math.*, 83 (1950), 101–129.
- [3] L. Sario and K. Oikawa, Capacity functions, Springer, 1969.
- [4] L. Sario and M. Nakai, Classification theory of Riemann surfaces, Springer, 1970.
- [5] J. Garnette, Analytic capacity and measure, Lecture notes in math. 297, Springer, 1970.
- [6] L. Zalcman, Analytic capacity and rational approximation, Lecture notes in math. 50, Springer, 1968.
- [7] T. W. Gamelin, Uniform algebras, Prentice-Hall, 1969.

170 (IX.2) Fundamental Groups

A \dagger continuous mapping f from the interval $I = \{t | 0 \leq t \leq 1\}$ into a topological space Y is called a **path** connecting the **initial point** $f(0)$ and the **terminal point** $f(1)$. In particular, a path satisfying $f(0) = f(1) = y_0$ is called a **loop** (or **closed path**) with y_0 as the **base point**. For a path f , the **inverse path** \bar{f} of f is defined by $\bar{f}(t) = f(1 - t)$. When the terminal point of f and the initial point of g coincide, the path F defined by $F(t) = f(2t)$ for $0 \leq t \leq 1/2$ and $F(t) = g(2t - 1)$ for $1/2 \leq t \leq 1$ is called the **product** (or **concatenation**) of f and g , and is denoted by $f \cdot g$. With $[f]$ standing for the equivalence class of a path f under the relation of \dagger homotopy relative to 0, 1 ($\in I$) (i.e., by homotopy with 0 and 1 being fixed), the inverse $[f]^{-1} = [\bar{f}]$ and the product $[f] \cdot [g] = [f \cdot g]$ are defined. In particular, in the set of homotopy

classes of loops with one common base point y_0 , the product is always defined, and the set forms a group $\pi_1(Y, y_0)$. This group is called the **fundamental group** (or **Poincaré group**) (H. Poincaré, 1985) of Y (with respect to y_0). If Y is \dagger arcwise connected, then $\pi_1(Y, y_0) \cong \pi_1(Y, y_1)$ for an arbitrary pair of points y_0, y_1 , and the structure of the group is independent of the choice of the base point. This group is denoted simply by $\pi_1(Y)$. A continuous mapping $\varphi: (Y, y_0) \rightarrow (Y', y'_0)$ induces a homomorphism $\varphi_*: \pi_1(Y, y_0) \rightarrow \pi_1(Y', y'_0)$ by sending $[f]$ to $\varphi_*[f] = [\varphi \circ f]$, and $(\varphi' \circ \varphi)_* = \varphi'_* \circ \varphi_*$ holds for the composite $\varphi' \circ \varphi$ of mappings. Thus $\pi_1(Y)$ is a \dagger topological invariant of Y . If $\pi_1(Y)$ consists of only one class (the class of the constant path), we say that Y is **simply connected**. For example, cells and spheres S^n ($n \geq 2$) are simply connected. The famous \dagger Poincaré conjecture states that a simply connected 3-dimensional compact \dagger manifold is homeomorphic to the 3-dimensional sphere. We have **van Kampen's theorem**: Let P be a connected polyhedron, P_1 and P_2 be its connected subpolyhedra such that $P_1 \cap P_2$ is connected, and $P = P_1 \cup P_2$. Then $\pi_1(P)$ is isomorphic to the group (\dagger amalgamated product) obtained from the \dagger free product of $\pi_1(P_1)$ and $\pi_1(P_2)$ by giving the relations that the images of each element of $\pi_1(P_1 \cap P_2)$ in $\pi_1(P_1)$ and in $\pi_1(P_2)$ are equivalent. Also, the fundamental group of the \dagger product spaces is the direct product of the fundamental groups of the spaces involved. Any group is the fundamental group of some \dagger CW complex. The Abelianization $\pi_1/[\pi_1, \pi_1]$ of the fundamental group $\pi_1 = \pi_1(Y)$ (Y arcwise connected) is isomorphic to the 1-dimensional integral \dagger homology group $H_1(Y)$. For example, the fundamental groups of a circle S^1 and a \dagger torus T^n are an infinite cyclic group and a free Abelian group of rank n , respectively; the fundamental group of a 1-dimensional CW complex is a free group; and the fundamental group of an orientable 2-dimensional closed surface of \dagger genus p is a group having $2p$ generators $\{a_1, \dots, a_p, b_1, \dots, b_p\}$ and a relation $a_1 b_1 a_1^{-1} b_1^{-1} \dots a_p b_p a_p^{-1} b_p^{-1} = 1$. If x_0 is a fixed point of the circle S^1 , then the fundamental group can be defined as the set of all homotopy classes of continuous mappings $f: (S^1, x_0) \rightarrow (Y, y_0)$.

Extending the definition of the fundamental group by replacing I, S^1 with I^n, S^n , we obtain the n -dimensional homotopy group (\rightarrow 202 Homotopy Theory; 91 Covering Spaces).

References

- [1] H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Teubner, 1934 (Chelsea, 1965).

[2] S. Lefschetz, Introduction to topology, Princeton Univ. Press, 1949.
[3] W. S. Massey, Algebraic topology, an introduction, Springer, 1977.
[4] E. H. Spanier, Algebraic topology, McGraw-Hill, 1966.

G

171 (XXI.25) Galois, Evariste

Evariste Galois (October 25, 1811–May 31, 1832) was born in Bourg-la-Reine, a suburb of Paris. In 1828, while still in junior high school, he published a paper on periodic continued fractions. Although he published four papers, his most important works were submitted to the French Academy of Science and either lost or rejected. He was unsuccessful in his attempt to enter the Ecole Polytechnique and instead entered the Ecole Normale Supérieure in 1829. Active in political affairs, he was expelled from school, imprisoned, and died in a duel soon after his release.

The night before the duel, he left his research outline and manuscripts to his friend, A. Chevalier. These were published by J. Liouville in *J. Math. Pures Appl.*, first series, 11 (1846). The contents include the concept of groups and what essentially became the †Galois theory of algebraic equations. The manuscript also contained expressions such as “theory of ambiguity,” which seems to indicate that Galois intended to study the theory of algebraic functions along the same lines.

References

- [1] E. Galois, *Oeuvres mathématiques*, E. Picard (ed.), Gauthier-Villars, 1897.
- [2] F. Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert I*, Springer, 1926 (Chelsea, 1956).
- [3] R. Bourgne and J.-P. Azra, *Ecrits et mémoires mathématiques d'Evariste Galois*, Gauthier-Villars, 1962.

172 (III.7) Galois Theory

A. History

After the discovery of formulas giving the general solutions of algebraic equations of degrees 3 and 4 in the 16th century, efforts to solve equations of degree 5 remained unsuccessful. Early in the 19th century P. Ruffini and N. H. †Abel showed that a general algebraic solution is impossible. Shortly afterward, E. †Galois established a general principle concerning the construction of roots of algebraic equations by radicals. The principle was described in terms of the structure of a certain permutation group (the Galois group)

of the roots of the equation. Even in this original form of the theory (**Galois theory**), Galois not only completed the research started by J. L. Lagrange, P. Ruffini, and Abel, but also made an epochal discovery that opened the way to modern algebra. J. W. R. Dedekind (*Werke* III, 1894) interpreted this result as a duality theorem concerning the automorphism groups of a field. It was shown later that Galois theory plays an important role in the general theory of commutative fields established by E. Steinitz. In the 1920s, W. Krull generalized the idea of Dedekind, using the concept of topological algebraic systems, and obtained the Galois theory of infinite algebraic extensions (*Math. Ann.*, 100 (1928)). The Galois theory gives a model for a successful theory summarizing the essentials of separable algebraic extensions and has led to analogous theories for other algebraic systems. For example, E. R. Kolchin constructed an analogous theory for differential fields where the Galois groups are algebraic groups (→ 113 Differential Rings, Galois theory of differential fields). Another line of development of this theory, also originated by Dedekind (*Werke* III, 1876/77), led to the Galois theory of rings, an object of active research by N. Jacobson (*Ann. Math.*, 41 (1940)), T. Nakayama, and others since the 1940s. Also, recently, the Galois theory for some general algebraic systems containing inseparable fields has been constructed by M. E. Sweedler, U. S. Chase, and others in which Galois groups are replaced by Hopf algebras or bialgebras (→ 203 Hopf algebras).

B. Definitions

Given a group G of †automorphisms of a given †field L , the subfield $F(G) = \{a \in L \mid a^\sigma = a, \sigma \in G\}$ is called the **invariant field** associated with G . For any extension fields L, L' of K , an isomorphism of L into L' whose restriction to K is the identity is called a **K -isomorphism**. If L is a †normal extension of K , any K -isomorphism of L into L' is a K -automorphism of L . If an algebraic extension L of K is normal, the group $G(L/K)$ of all K -automorphisms of L is called the **Galois group** of L/K . A †separable normal algebraic extension of K is called a **Galois extension** of K . Let L/K be a finite normal extension; then there exist intermediate fields M and N of L/K such that M/K is a Galois extension and N/K is a †purely inseparable extension and $L = M \otimes_K N$ (→ 277 Modules J). Further, $G(L/K) = G(L/N) \simeq G(M/K)$ and the order of $G(L/K)$ is equal to the †separable degree of L/K , i.e. $[M:K]$. A necessary and sufficient condition for L/K to

be a Galois extension is that the invariant field associated with $G(L/K)$ be K . A Galois extension L/K is called an **Abelian extension** or a **cyclic extension** when $G(L/K)$ is \dagger Abelian or \dagger cyclic, respectively (\rightarrow 149 Fields).

C. Fundamental Theorem of Galois Theory

Let L/K be a finite Galois extension and G its Galois group. Then there exists a \dagger dual lattice isomorphism between the set of intermediate fields of L/K and the set of subgroups of G , under which an intermediate field M of L/K corresponds to the subgroup $H = G(L/M)$; conversely, a subgroup H of G corresponds to $M = F(H)$. The degree of extension $[L:M]$ is equal to the order of the corresponding subgroup H (in particular, $[L:K]$ is the order of G), and $[M:K]$ coincides with the index $(G:H)$. If subfields M and M' are \dagger conjugate over K , then the corresponding subgroups $G(L/M)$ and $G(L/M')$ are conjugate to each other in G , and vice versa. In particular, M/K is a Galois extension if and only if the subgroup H corresponding to M is a \dagger normal subgroup of G , and in this case, the Galois group $G(M/K)$ is isomorphic to the factor group G/H .

D. Extensions of a Ground Field

Let L/K be a finite Galois extension, K'/K any extension, and L' the \dagger composite field of L and K' . Then L'/K' is also a Galois extension, and its Galois group is isomorphic to $G(L/L \cap K')$ by the restriction map.

E. Normal Basis Theorem

Let L/K be a finite Galois extension with Galois group G . Then there exists an element u of L such that $\{u^\sigma \mid \sigma \in G\}$ forms a basis for L over K called a **normal basis**. If we denote by $K[G]$ the \dagger group ring of G over K , a $\dagger K[G]$ -module structure can be introduced in L by the operation $\sum a_\sigma \sigma(x) = \sum a_\sigma x^\sigma$; the existence of a normal basis implies that L is isomorphic to $K[G]$ itself as a $K[G]$ -module, or in other words, that the K -linear representation of G by means of L is equivalent to the \dagger regular representation of G .

F. Examples of Galois Extensions

(1) **\dagger Cyclotomic Fields.** Let m be a positive integer not divisible by the \dagger characteristic of K ; ζ a \dagger primitive m th root of unity, and $L = K(\zeta)$. Then L/K is an Abelian extension, and its Galois group is isomorphic to a subgroup of

the \dagger reduced residue class group $(\mathbb{Z}/m\mathbb{Z})^*$; in particular, if $K = \mathbb{Q}$, the subgroup coincides with $(\mathbb{Z}/m\mathbb{Z})^*$, by the irreducibility of \dagger cyclotomic polynomials. Hence the degree $[Q(\zeta):Q]$ is equal to $\varphi(m)$, where φ is \dagger Euler's function.

(2) **Finite Fields.** A \dagger finite field K has nonzero characteristic p , and the number q of elements of K is a power of p . Also, K is uniquely determined by q (up to isomorphism), hence it is denoted by $GF(q)$ or F_q . Thus $GF(q^n)$ is the only extension of $GF(q)$ of degree n ; moreover, it is a cyclic extension.

(3) **Kummer Extensions.** Assume that K contains a primitive m th root ζ of unity and the characteristic of K is 0 or is not a divisor of m . Denote by K^* the multiplicative group of K . An extension L of K can be expressed in the form $L = K(\sqrt[m]{a_1}, \dots, \sqrt[m]{a_r})$ ($a_i \in K$) if and only if L/K is an Abelian extension and all $\sigma \in G(L/K)$ satisfy $\sigma^m = 1$; in this case L/K is called a **Kummer extension of exponent m** . There exists a one-to-one correspondence between Kummer extension L of exponent m over K and finite subgroups $H/(K^*)^m$ of the factor group $K^*/(K^*)^m$, given by the relations $H = L^m \cap K^*$, $L = K(\sqrt[m]{H})$. Moreover, there exists a canonical isomorphism between $H/(K^*)^m$ and the \dagger character group of $G(L/K)$, so that $H/(K^*)^m$ is isomorphic to $G(L/K)$. Let $L = K(\theta)$ be a cyclic Kummer extension of degree m of K , and let σ be a generator of the Galois group $G(L/K)$. Then the **Lagrange resolvent** $(\zeta, \theta) = \theta + \zeta\theta^\sigma + \dots + \zeta^{m-1}\theta^{\sigma^{m-1}}$ satisfies $(\zeta, \theta)^\sigma = \zeta^{-1}(\zeta, \theta)$, $(\zeta, \theta)^m \in K$, and θ and its conjugates can be expressed in terms of (ζ, θ) . In particular, L is generated by (ζ, θ) over K .

(4) **Artin-Schreier Extensions.** Assume that K is of characteristic $p \neq 0$. For any element a of an extension of K , we denote by $\mathcal{P}a$ the element $a^p - a$ and by $(1/\mathcal{P})a$ a root of $\mathcal{P}X - a = 0$. A finite extension L of K is of the form $L = K((1/\mathcal{P})a_1, \dots, (1/\mathcal{P})a_r)$ ($a_i \in K$) if and only if L/K is a Galois extension whose Galois group is an Abelian group of \dagger exponent p ; in this case, L/K is called an **Artin-Schreier extension**. There exists a one-to-one correspondence between Artin-Schreier extensions L over K and finite subgroups $H/\mathcal{P}K$ of the additive group $K/\mathcal{P}K$, given by the relations $H = \mathcal{P}L \cap K$, $L = K((1/\mathcal{P})H)$; moreover, $H/\mathcal{P}K$ is isomorphic to the character group of $G(L/K)$ (therefore also to $G(L/K)$ itself). More generally, for Abelian extensions L of **exponent p^n** (i.e., Galois extensions whose Galois groups are Abelian groups of exponent p^n), we obtain similar descriptions by using the additive

group of \dagger Witt vectors of length n instead of K (\rightarrow 449 Witt Vectors).

G. Galois Group of an Equation

L/K is a finite Galois extension if and only if L is a \dagger minimal splitting field of a \dagger separable polynomial $f(X)$ in $K[X]$. In this case, we call $G(L/K)$ the **Galois group of the polynomial** $f(X)$ or of the **algebraic equation** $f(X)=0$. The equation $f(X)=0$ is called an **Abelian equation** or a **cyclic equation** if its Galois group is Abelian or cyclic, respectively, while $f(X)=0$ is called a **Galois equation** if L is generated by any root of $f(X)$ over K . Generally, $G(L/K)$ can be \dagger faithfully represented as a permutation group of roots of $f(X)=0$. If this group is \dagger primitive, then $f(X)=0$ is called a **primitive equation**. The index of the group in the group of all permutations of roots is called the **affect** of the equation $f(X)=0$; if the affect is 1, the equation $f(X)=0$ is called **affect-less**. Let u_1, \dots, u_n be \dagger algebraically independent elements over K . Then for the polynomial $F_n(X)=X^n-u_1X^{n-1}+\dots+(-1)^nu_n$ in $K(u_1, \dots, u_n)[X]$, the equation $F_n(X)=0$ is called a **general equation** of degree n . The Galois group of $F_n(X)=0$ is isomorphic to the \dagger symmetric group \mathfrak{S}_n of degree n , and if K is not of characteristic 2, then the quadratic subfield corresponding to the \dagger alternating group \mathfrak{A}_n is the field $K(\sqrt{D})$ obtained by adjoining the quadratic root of the \dagger discriminant D of $F_n(X)$.

H. Solvability of an Algebraic Equation

Assume that K is of characteristic 0, $f(X) \in K[X]$, and L is the minimal splitting field of $f(X)$. We say that the equation $f(X)=0$ is **solvable by radicals** if there is a chain of subfields $K=L_0 \subset L_1 \subset \dots \subset L_r=L$ such that $L_i=L_{i-1}(\sqrt[n_i]{a_i})$ with some $a_i \in L_{i-1}$, and this is the case if and only if the Galois group of $f(X)$ is \dagger solvable (Galois). In particular, Abelian equations are solvable by radicals. Cyclic equations are solved by using the Lagrange resolvent, and theoretically the general solvable equation can be solved by repeating this procedure. Since \mathfrak{S}_n is solvable if and only if $n \leq 4$, it follows that a general equation of degree n is solvable only if $n=1, 2, 3, 4$ (Abel). (For a method of solving these equations \rightarrow 10 Algebraic Equations D.) Also, a polynomial is solvable by square roots if and only if the order of the Galois group is a power of 2. This fact enables us to answer some questions concerning geometric construction problems such as trisection of an angle or division

of a circumference in equal parts (\rightarrow 179 Geometric Construction).

I. Infinite Galois Extensions

If a Galois extension L/K is infinite, then its Galois group is an infinite group. Let $\{M_v\}$ be the family of intermediate fields of L/K that are finite and normal over K , and put $H_v = G(L/M_v)$. Then by taking $\{H_v\}$ as a \dagger base of a neighborhood system of the unity element, G becomes a \dagger topological group. This topology is called the **Krull topology** (Krull, *Math. Ann.*, 100 (1928)). G is then isomorphic to the \dagger projective limit of the family of finite groups $\{G/H\}$ and is \dagger totally disconnected and \dagger compact. There is a one-to-one correspondence between the set of intermediate fields of L/K and the set of closed subgroups of G given by the map (Galois group) \leftrightarrow (invariant field), and thus we have a generalization of Galois theory for finite extensions as described in Section C. Various theories, including the theory of Kummer extensions, can be generalized to the case of infinite extensions (\rightarrow 423 Topological Groups).

J. Galois Cohomology

Let L/K be a finite Galois extension and G its Galois group. Then both the additive group L and the multiplicative group L^* have G -module structures. The \dagger cohomology groups of G with coefficient module L are 0 for all dimensions because of the existence of a normal basis (\rightarrow 200 Homological Algebra). As for the multiplicative group L^* , we have $\hat{H}^0(G, L^*) \cong K^*/N(L^*)$ (N is the \dagger norm $N_{L/K}$), $H^1(G, L^*)=0$ (**Hilbert's theorem 90** or the **Hilbert-Speiser theorem**). In particular, if G is a cyclic group with generator σ , then every element a such that $N(a)=1$ can be expressed in the form $a=b^{1-\sigma}$. $H^2(G, L^*)$ is isomorphic to the \dagger Brauer group of \dagger central simple algebras over K which have L as a \dagger splitting field. In the case of number fields, a number of G -modules arise, such as \dagger principal orders, \dagger unit groups, \dagger ideal groups, \dagger idele groups, and so on, whose cohomological considerations are important (\rightarrow 6 Adeles and Ideles; 59 Class Field Theory). Further, let A be a group (not necessarily commutative), and suppose that G acts on A . We denote by ${}^\sigma a$ the action of $\sigma \in G$ on $a \in A$. A is called a **G -group** if ${}^\sigma(ab) = {}^\sigma a {}^\sigma b$ for all $a, b \in A$ and $\sigma \in G$. Then, in the same way as for G -modules, we can define the **0th cohomology group** $H^0(G, A)$ to be the subgroup A^G of A consisting of all elements of A left fixed by G . The map $a: \sigma \mapsto a_\sigma$ from G into A such that $a_{\sigma\tau} = a_\sigma {}^\sigma a_\tau$ ($\sigma, \tau \in G$) is called a **1-cocycle** with

values in A . Denote by $Z^1(G, A)$ the set of the 1-cocycles of G with values in A . Two 1-cocycles a and a' in $Z^1(G, A)$ are called **cohomologous** if there exists an element $b \in A$ such that $a'_\sigma = b^{-1} a_\sigma b$ for all $\sigma \in G$. This is an equivalence relation in $Z^1(G, A)$ and the 1-cohomology set $H^1(G, A)$ of G with values in A is defined to be the set of cohomologous classes of $Z^1(G, A)$. $H^1(G, A)$ does not have the structure of a group in general, but there does exist an element called an identity, namely, the cohomologous class containing the trivial 1-cocycle. In general, a set X with an element p of X is called a **pointed set**; for any two pointed sets (X, p) and (Y, q) , a map $f: X \rightarrow Y$ is called a **morphism of pointed sets** if $f(p) = q$. For pointed sets (X, p) , (Y, q) , and (Z, r) , a sequence of morphisms $X \xrightarrow{f} Y \xrightarrow{g} Z$ of pointed sets is called **exact** if $\text{Im } f = \text{Ker } g$. The 1-cohomology set $H^1(G, A)$ with the identity can be regarded as a pointed set. Therefore, exact sequences of these sets make sense and possess some of the properties of cohomology groups in the commutative case. For example, let $1 \rightarrow A \rightarrow B \rightarrow C \rightarrow 1$ be an exact sequence of G -groups (and A is central in B); then

$$H^0(G, A) \rightarrow H^0(G, B) \rightarrow H^0(G, C) \\ \rightarrow H^1(G, A) \rightarrow H^1(G, B) \rightarrow H^1(G, C) (\rightarrow H^2(G, A))$$

is an exact sequence of pointed sets. For a Galois extension L/K with Galois group G , a linear algebraic L -group defined over K (\rightarrow 13 Algebraic Groups) has naturally the structure of a G -group, and we have $H^1(G, GL_n(L)) = 0$. Applying the above exact sequence to $1 \rightarrow SL_n(L) \rightarrow GL_n(L) \xrightarrow{\det} L^* \rightarrow 1$, we have $H^1(G, SL_n(L)) = 0$. Also, we have $H^1(G, Sp_{2n}(L)) = 0$ for any integer $n \geq 1$. It is difficult to define higher cohomology sets naturally, but various methods to define them have been obtained.

In many cases, we are more concerned with the \dagger category of Galois extensions of K with K -isomorphisms between them than with a single extension L/K . In other words, we consider a \dagger functor $L \rightarrow \mathfrak{F}(L)$ of the category of Galois extensions of K into the category of (Abelian) groups, and study the cohomology related to $G(L/K)$ -module or $G(L/K)$ -group structures derived from $\mathfrak{F}(L)$. In the case of infinite algebraic extensions, we consider the inductive limit of cohomology of subfields of finite degrees, making use of continuous cocycles of Galois groups relative to the Krull topology [8, 9].

Let L/K be a Galois extension with Galois group G . Consider objects X, Y defined over K on which the extension of the ground field is defined (such as K -linear spaces with certain tensors on them, algebras over K (associative

or not), K -varieties, or K -algebraic groups defined over K). When X and Y are isomorphic over L , Y is called a L/K -form of X . Let $E(L/K, X)$ be the set of all isomorphism classes of X over L . Let A be the group of all automorphisms over L of X . For a L/K -form Y of X and an isomorphism $f: X \rightarrow Y$ over L , since G acts on X and Y , one can define the isomorphism ${}^a f: X \rightarrow Y$ over L that satisfies ${}^a(f(x)) = {}^a f({}^a x)$. In particular, A is a G -group by this action. Further, for $\sigma \in G$, define $a_\sigma = f^{-1} \cdot {}^\sigma f \in A$, then $a: \sigma \mapsto a_\sigma$ is a 1-cocycle of G with values in A that corresponds to a L/K -form Y of X . This induces a bijection from $E(L/K, X)$ onto $H^1(G, A)$. Thus L/K -forms of X can be classified if one can determine $H^1(G, A)$. For example, $H^1(G, Sp_{2n}(L)) = 0$ is equivalent to saying that an L/K -form of a skew-symmetric bilinear form on a K -linear space of dimension $2n$ is unique up to K -isomorphisms. L/K -forms of a semisimple Lie algebra \mathfrak{g} over K can be classified by the 1-cohomology set of the algebraic group $\text{Aut}(\mathfrak{g} \otimes_K L)$. (For the L/K -forms of algebraic groups \rightarrow 13 Algebraic Groups M.) This descent theory can also be discussed for more general categories.

K. Galois Theory of Rings

The theory of \dagger centralizers in simple algebras can be interpreted as the theory of a certain Galois correspondence with respect to inner automorphism groups. Also, by using \dagger crossed products, we can deduce from the theory of centralizers in simple algebras the Galois theory of commutative fields. On the other hand, Jacobson obtained a Galois theory of \dagger division rings with respect to finite groups of outer automorphisms that is similar to the commutative case. Since then, many algebraists have proceeded with investigations that aim either at unifying these two theories by admitting inner automorphisms in the group of automorphisms, at extending the theory from division rings to general rings such as simple rings, \dagger primitive rings, or \dagger semiprimary rings, or at weakening the finiteness conditions. One principal method in these theories lies in considering first the endomorphism ring $\text{Hom}_R(S, S)$ for an extension S/R and then the roles of endomorphisms (or derivations) in it [6] (\rightarrow 29 Associative Algebras).

Let K be a field of characteristic $p > 0$ and L/K be a finite purely inseparable extension such that $L^p \subset K$. The set of all \dagger derivations of L/K forms a restricted Lie algebra $D(L/K)$ over K . Then there exists a one-to-one and dual lattice isomorphic correspondence between the set of intermediate fields M of L/K

and the set of restricted Lie subalgebras H of $D(L/K)$, given by the relations $H = D(L/M)$ and $M = \{a \in L \mid d(a) = 0, d \in H\}$ (Jacobson [7]). Namely, in the case of purely inseparable extensions, derivations or higher derivations play the role of automorphisms, and the bialgebras defined by such derivations correspond to Galois groups of Galois extension. Now, the group algebra KG of a group G over K and the (restricted) universal enveloping algebra of a (restricted) Lie algebra over K both have the structure of a * Hopf algebra. From this fact, unifying the Galois theory of Galois extensions and Jacobson's theory for purely inseparable extensions and using bialgebras or Hopf algebras, one can construct Galois theories of more general objects containing certain nonseparable field extensions (see M. E. Sweedler, *Ann. Math.*, 87 and 88 (1968); S. U. Chase and Sweedler, *Lecture notes in math.* 97, Springer, 1969).

References

- [1] J. W. R. Dedekind, Über die Permutationen des Körpers aller algebraischen Zahlen, *Gesammelte mathematische Werke*, Viewig, 1930–1932, vol. 2, 272–291.
- [2] E. Artin, *Galois theory*, Univ. of Notre Dame Press, second edition, 1948.
- [3] N. Bourbaki, *Eléments de mathématique*, *Algèbre*, ch. 5, *Actualités Sci. Ind.* 1102b, Hermann, second edition, 1959.
- [4] N. G. Chebotarev, *Grundzüge der Galois'schen Theorie*, Noordhoff, 1950.
- [5] B. L. van der Waerden, *Algebra I*, Springer, seventh edition, 1966.
- [6] N. Jacobson, *Structure of rings*, Amer. Math. Soc. Colloq. Publ., 1968.
- [7] N. Jacobson, *Lectures in abstract algebra III*, Van Nostrand, 1964; Springer, 1975.
- [8] J.-P. Serre, *Corps locaux*, *Actualités Sci. Ind.*, Hermann, 1962; *Local fields*, Springer, 1979.
- [9] J.-P. Serre, *Cohomologie galoisienne*, *Lectures notes in math.* 5, Springer, 1974.

173 (XIX.9) Game Theory

A. Introduction and Historical Highlights

Game theory consists of mathematical models used in the study of decision making in situations involving conflict and cooperation. A conflict arises when each player in a game

selects from a list of alternatives one which, possibly together with chance and random events, leads to various outcomes over which the players have different preferences; thus the behavior of one player aiming at his own favorable outcome might induce unfavorable outcomes for others. Although the potential outcomes usually bring about conflicts among the players, there may be room for cooperation among some of them. Game theory attempts to extract that which is common and essential to such situations, to handle them by means of mathematical methods, and to provide a normative guide to rational behavior for each of the players. Game theory thus goes beyond classical theories of probability and decision making that are sufficient to solve games involving just one player and chance.

Modern game theory started in 1944 with the publication of the monumental book by von Neumann and Morgenstern [1]. These authors presented many logical classifications of games, including a distinction between two-person and n -person games, between **constant-sum** (zero-sum) and **general-sum** games (depending upon whether the sum of payoffs to the players is constant (zero) or not), and between **noncooperative** and **cooperative** games (depending upon whether any collaboration among the players is prohibited or allowed). The second edition included a remarkable expected utility theory, which has become a mainstay of game theory. Von Neumann and Morgenstern also gave three representations of games. The first representation is the so-called extensive form; this representation has been slightly modified by Kuhn [2]. The second representation is the normal-form game; this is the form which the minimax theorem for two-person zero-sum games was established [3]. This theorem was generalized to n -person general-sum noncooperative games by Nash [4]. Finally, the characteristic-function form was the one in which the authors developed the theory of stable sets. Several other solution concepts, such as the core, the Shapley value, and the bargaining set, have since been defined for games in characteristic-function form. Historically, the development of game theory has been closely related to various areas of pure mathematics, such as analysis, topology, geometry, and the foundations of mathematics. A survey of game theory up to 1957 was presented by Luce and Raiffa [5]. In the 1950s and early 1960s, several major papers appeared in five issues of the *Annals of Mathematics Studies* [6]. Lucas [7] has presented a good survey of developments to 1972, and Shubik [8] has surveyed the development of the field through 1981. Current articles

appear mainly in the *International Journal of Game Theory*, and also in the journals in fields such as operations research, management science, economics, political science, and psychology.

B. The Extensive Form

An *n*-person game in extensive form is represented by a game †tree (i.e., a connected graph with no cycles (→ 186 Graph Theory)) having the following properties: There is one special vertex corresponding to the starting point of the game. Each nonterminal vertex corresponds either to a **move** of one of the *n* players or to a **chance move**. The edges ascending from a vertex denote the **alternatives** of the player at this vertex. For each terminal vertex there is an *n*-dimensional vector whose components represent the **payoffs** to each player. The state of a player's information at any stage can be described by certain subsets of the set of all his vertices, called **information sets**. At each of his moves, player *i* knows which information set he is in, but not which vertex he occupies within this set. A **local strategy** for player *i* is a †probability distribution over the set of all alternatives at each of his information sets. A **behavior strategy** for player *i* is a function which assigns a local strategy to all of his information sets. A **pure strategy** is a special behavior strategy that assigns a particular choice to each information set. Kuhn [2] showed the existence of pure optimal strategies for *n*-person general-sum noncooperative games with perfect information (i.e., games in which all information sets contain a single vertex) as a generalization of the result for two-person zero-sum games given in [1]). Kuhn also proved the existence of equilibrium behavior strategies for games with perfect recall (i.e., games in which player *i* at any of his information sets remembers all his prior moves but is not aware of the prior choices of the other players). The definition of "equilibrium strategy" is given in the next section. Because of the complexity of *n*-person general-sum games, research on them has been minimal since the work of Kuhn. Recently a new attack on them has begun; for example, the Nash equilibrium point in extensive forms has been reexamined by Selten [9].

C. The Normal Form (or Strategic Form)

An *n*-person game in normal form is specified by $\{N, \{X^i\}_{i \in N}, \{F^i\}_{i \in N}\}$, where $N = \{1, \dots, n\}$ is a set of players, X^i is the set of player *i*'s strategies, and F^i is a real-valued function on

the Cartesian product $\prod_{i=1}^n X^i$, called player *i*'s **payoff function**. Chance is incorporated by invoking an extra set X^0 of chance moves and a probability distribution on X^0 . A two-person zero-sum game or **matrix game** is the simplest case in which the existence of an equilibrium point has been established. Equilibrium follows from von Neumann's **minimax theorem**: Let $M^1 = \{1, \dots, m_1\}$, $M^2 = \{1, \dots, m_2\}$ be the sets of pure strategies that may be chosen by two players. Let a_{ij} be player 1's payoff when strategies *i* and *j* are taken by players 1 and 2. A **mixed strategy** for player *i* is a probability distribution on M^i . Sets of mixed strategies for players 1 and 2 are thus given by $X^1 = \{x \in R^{m_1} | \sum_{i=1}^{m_1} x_i = 1, x_i \geq 0 \forall i \in M^1\}$ and $X^2 = \{x \in R^{m_2} | \sum_{i=1}^{m_2} x_i = 1, x_i \geq 0 \forall i \in M^2\}$. If players 1 and 2 use the mixed strategies x^1 and x^2 , respectively, the expected payoff for player 1 is $F^1(x^1, x^2) = \sum_i \sum_j x_i^1 a_{ij} x_j^2$, which is a payoff function for player 1. Von Neumann [3] proved

$$\max_{x^1 \in X^1} \min_{x^2 \in X^2} F^1(x^1, x^2) = \min_{x^2 \in X^2} \max_{x^1 \in X^1} F^1(x^1, x^2).$$

We say that a pair (\hat{x}^1, \hat{x}^2) satisfying the foregoing minimax theorem is an **equilibrium point**. This pair (\hat{x}^1, \hat{x}^2) turns out to be a saddle point (→ 292 Nonlinear Programming A) of $F^1(x^1, x^2)$. The duality theorem (→ 255 Linear Programming B) is mathematically equivalent to this minimax theorem. A generalization of this equilibrium to *n*-person general-sum noncooperative games was presented by Nash [4]. An *n*-tuple $(\hat{x}^1, \dots, \hat{x}^n)$ ($\hat{x}^i \in X^i$) is a **Nash equilibrium** if, for each *i*,

$$F^i(\hat{x}^1, \dots, \hat{x}^{i-1}, \hat{x}^i, \hat{x}^{i+1}, \dots, \hat{x}^n) \geq F^i(\hat{x}^1, \dots, \hat{x}^{i-1}, x^i, \hat{x}^{i+1}, \dots, \hat{x}^n) \text{ for all } x^i \in X^i.$$

That is, no player can improve his payoff by changing his strategy if all other players continue to use the same strategies. Nash demonstrated the existence of this equilibrium for *n*-person general-sum noncooperative games with finite pure strategies. The existence of Nash equilibria for wider classes of noncooperative games is proved by means of fixed-point theorems (→ 153 Fixed-Point Theorems). Recently, **multistage games**, such as supergames and stochastic games in which games are played repeatedly, have become a major research topic of noncooperative game theory (with work being done on both extensive and normal forms). In two-person general-sum games or **bimatrix games**, there arise possibilities of cooperation (or bargaining) between two players. A solution concept for such situations, the **Nash bargaining solution**, was proposed by Nash [10]; this solution was subsequently discussed by Luce and Raiffa [5].

D. The Characteristic-Function Form (or Coalitional Form)

An n -person cooperative game in characteristic-function form is given by a pair (N, v) , where $N = \{1, \dots, n\}$ is a set of players and v is a real-valued function on a set of all subsets of N with $v(\emptyset) = 0$, called a **characteristic function**. Sometimes v is assumed to be **superadditive**, i.e., $v(S \cup T) \geq v(S) + v(T) \forall S, T \subseteq N$ with $S \cap T \neq \emptyset$. $v(S)$ represents the worth or the power achievable by the subset (coalition) S when its members cooperate regardless of the behavior(s) of the players in the complement of S . An n -dimensional vector $x = (x_1, \dots, x_n)$ is said to be an **imputation** if it satisfies (i) $x_i \geq v(\{i\}) \forall i \in N$ (individual rationality) and (ii) $\sum_{i \in N} x_i = v(N)$ (group rationality). The set of all imputations, denoted by A , represents all reasonable or realizable ways of distributing the available gains among the n players. An imputation x is said to **dominate** another imputation y if there is some nonempty coalition S such that (i) $x_i > y_i \forall i \in S$ and (ii) $\sum_{i \in S} x_i \leq v(S)$ (the effectiveness of S with respect to x). The original solution concept for games in characteristic-function form given by von Neumann and Morgenstern [1] is now called the **stable-set** or **von Neumann–Morgenstern solution**. A subset K of A is a stable set if (i) no dominance relation exists between any two elements of K (internal stability) and (ii) any imputation outside K is dominated by some imputation of K (external stability). The existence of stable sets was settled negatively by the ten-person example of Lucas [11]. This example, however, is rather specialized, and thus the existence of stable sets may yet be proved for a large class of games. Stable sets have been characterized for several classes of games, and these sets accurately reflect the coalition-forming processes among the players. Thus stable-set theory remains a major research topic in the theory of games in characteristic-function form. Many other solution concepts have been developed since that of the stable set. One of these is the **core**, defined by Gillies in [6] (its naive idea had already appeared in [1]). The core is the set $C = \{x \in A \mid \sum_{i \in S} x_i \geq v(S) \forall S \subseteq N\}$. This says that no coalition can protest against or block an imputation in the core on grounds that the coalition can expect more. For superadditive games the core coincides with the set of undominated imputations, and thus the core is a subset of any stable set if both exist. The condition for nonemptiness of the core was derived by Shapley [12] using the duality theorem. Another solution concept, defined by Shapley in [6], is known as the **Shapley value**. The Shapley value is a function $\varphi(v) = (\varphi_1(v),$

$\dots, \varphi_n(v))$ sending an arbitrary characteristic function v on N into an n -dimensional Euclidean space satisfying the conditions (i) $\varphi_{\pi i}(\pi v) = \varphi_i(v)$, where π is any permutation of N and $\pi v(\pi S) = v(S) \forall S \subseteq N$; (ii) $\sum_{i \in S} \varphi_i(v) = v(S) \forall S \subseteq N$ such that $v(T) = v(S \cap T) \forall T \subseteq N$; and (iii) $\varphi_i(v + w) = \varphi_i(v) + \varphi_i(w) \forall i \in N$ and for any two games v and w . These three axioms uniquely determine the value

$$\varphi_i(v) = \sum_{\substack{S \subseteq N \\ S \ni i}} \frac{(s-1)!(n-s)!}{n!} (v(S) - v(S - \{i\}))$$

for each i , where s is the number of players in S . The Shapley value is an imputation, and can be viewed as a fair-division solution since the three axioms are desirable properties for any equitable allocation scheme. From the above formula, the Shapley value can also be interpreted as the average of the marginal contributions of the players in a coalition. A **bargaining-set** concept was proposed by Aumann and Maschler in [6]. This set describes what payoffs are stable once a particular **coalition structure** (a \dagger partition of N) has formed. Briefly, a payoff associated with a coalition structure is stable or in a bargaining set if there is no objection to it from any player; or, even with an objection, if there exists a counterobjection to such an objection from other players. For details \rightarrow [6]. Since there are many different ways to define objection and counterobjection, there are various types of bargaining sets. Some of them are known to be nonempty. Two additional solution concepts, the **kernel** and the **nucleolus**, derive from investigations into particular bargaining sets. The kernel, introduced by Davis and Maschler [13], is always a nonempty subset of its bargaining set. More important is the nucleolus defined by Schmeidler [14], which is a unique imputation in the kernel and thus in the bargaining set. It is also in the core if the latter is nonempty. The “excess” of any nonempty $S \subseteq N$ for an imputation x is defined by $e(x, S) = v(S) - \sum_{i \in S} x_i$. The excess represents the dissatisfaction or the complaint of the coalition S with respect to x . The nucleolus is the imputation which minimizes the largest excess. If we have a tie, i.e., if the maximum excess attains a minimum at several imputations, then the next largest excess is to be compared, and so on. That is, the nucleolus is the \dagger lexicographical minimum in the ordering of these arrangements. It can be computed by solving a series of linear programs.

Several generalizations and variations of the classical von Neumann–Morgenstern formulation of games in characteristic-function form have also been investigated, such as the

games without side payments studied by Aumann and others (→ e.g., M. Shubik (ed.), *Essays in Mathematical Economics: In honor of Oskar Morgenstern*, Princeton Univ. Press, 1967); the **games in partition-function form** proposed by Thrall and Lucas (R. M. Thrall, and W. F. Lucas, *Naval Res. Logistics Quart.*, 10 (1963), 281–298); and the **games with infinitely many players**, in connection with which the Shapley value theory has become a major research topic [18].

E. Applications; Related Areas of Mathematics

Game theory has been applied to many fields, such as economics, political science, management science, operations research, information theory, and control theory, as well as to pure mathematics. Games in extensive form are now important tools for analyzing the effects of information, and thus for solving many decision problems with uncertainty. Non-cooperative games in normal form and Nash equilibria have been used in the study of many phenomena, including oligopolistic markets (Friedman [15]), bidding processes, electoral competition, resource allocation, and arms control. Cooperative games have been successfully applied to economics, and the relation between the core and competitive equilibria sheds further light on the theory of competitive economy. It is generally true that competitive equilibria are contained in the core. Debreu and Scarf [16] demonstrated that if the number of players approach infinity in a certain manner, the core shrinks to the set of competitive equilibria. By working in measure-theoretic terms, Aumann [17] was able to identify the core with the set of competitive equilibria. It was also demonstrated by Aumann and Shapley [18] that the set of competitive equilibria (and hence also the core) converges to the Shapley value under such formulations, provided certain conditions are satisfied. Another major application of cooperative game theory has arisen in political science, wherein value-type solutions, such as the Shapley value, are widely used as indices of the power of each participant in various voting situations. Major applications are to problems of cost allocation for public goods such as water resources [19], public transportation systems, and telephone systems; in such applications the core, the Shapley value, and the nucleolus have all been employed. The books [8, 20–22] are good references to the most recent applications of game theory.

Game theory also has many close relations with various areas of pure mathematics. The following are typical examples. The study of certain types of games in extensive form has

173 Ref. Game Theory

increased our understanding of the axiom of choice (→ 34 Axiom of Choice and Equivalents A) and other foundational questions. Nash equilibrium is closely related to fixed-point theorems (→ 153 Fixed-Point Theorems) and to separation theorems (→ 89 Convex Sets A). Cooperative game theory also has many connections to functional analysis and to convex analysis.

Finally, we mention that the study of differential games (→ 108 Differential Games) is also highly developed and widely used in areas such as economics and control theory.

References

- [1] J. von Neumann and O. Morgenstern, *Theory of games and economic behavior*, Princeton University Press, 1944; second edition, 1947; third edition, 1953.
- [2] H. W. Kuhn, Extensive games, *Proc. Nat. Acad. Sci. US*, 36 (1950), 570–576.
- [3] J. von Neumann, Zur Theorie der Gesellschaftsspiele, *Math. Ann.* 100 (1928), 295–320.
- [4] J. F. Nash, Noncooperative games, *Ann. Math.*, 54 (1951), 286–295.
- [5] R. Luce and H. Raiffa, *Games and decisions: Introduction and critical survey*, Wiley, 1957.
- [6] A. W. Tucker et al. (eds.), *Contributions to the theory of games*, vols. I–IV, and *Advanced studies in game theory*: *Ann. Math. Studies* 24, 28, 39, 40, 52, Princeton University Press, 1950, 1953, 1957, 1959, 1964.
- [7] W. F. Lucas, An overview of the mathematical theory of games, *Management Sci.*, 18, no. 5, pt. 2 (1972), P-3–P-19.
- [8] M. Shubik, *Game theory in the social sciences: Concepts and solutions*, MIT Press, 1982.
- [9] R. Selten, Reexamination of the perfectness concept for equilibrium points in extensive games, *Int. J. Game Theory*, 4 (1975), 25–55.
- [10] J. Nash, The bargaining problem, *Econometrica*, 18 (1950), 155–162.
- [11] W. F. Lucas, A game with no solution, *Bull. Amer. Math. Soc.*, 74 (1968), 237–239.
- [12] L. S. Shapley, On balanced sets and cores, *Naval Res. Logistics Quart.*, 14 (1967), 453–460.
- [13] M. D. Davis and M. Maschler, The kernel of a cooperative game, *Naval Res. Logistics Quart.*, 12 (1965), 223–259.
- [14] D. Schmeidler, The nucleolus of a characteristic function game, *SIAM J. Appl. Math.*, 17 (1969), 1163–1170.
- [15] J. W. Friedman, *Oligopoly and the theory of games*, North-Holland, 1977.
- [16] G. Debreu and H. Scarf, A limit theorem on the core of an economy, *Int. Econ. Rev.*, 4 (1963), 235–246.

- [17] R. J. Aumann, Markets with a continuum of traders, *Econometrica*, 32 (1964), 39–50.
 [18] R. J. Aumann and L. S. Shapley, Values of non-atomic games, Princeton Univ. Press, 1974.
 [19] M. Suzuki and M. Nakayama, The cost allocation of the cooperative resource development: A game-theoretical approach, *Management Sci.*, 22 (1976), 1081–1086.
 [20] P. C. Ordeshook (ed.), Game theory and political science, New York Univ. Press, 1978.
 [21] S. J. Brams et al. (eds.), Applied game theory, Physica-Verlag, 1979.
 [22] M. Shubik, A game-theoretic approach to political economy, MIT Press, 1984.

174 (XIV.4) Gamma Function

A. The Gamma Function

The function $\Gamma(x)$ was defined by L. Euler (1729) as the infinite product

$$\frac{1}{x} \prod_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)^x \left(1 + \frac{x}{n}\right)^{-1}.$$

Legendre later called it the **gamma function** or **Euler's integral of the second kind**. The latter name is based on the fact that for positive real x , we have

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt.$$

This function satisfies the functional relation

$$\Gamma(x+1) = x\Gamma(x),$$

and hence for positive integral x , we have $\Gamma(x+1) = x!$. C. F. Gauss denoted the function $\Gamma(x+1)$ by $\Pi(x)$ or $x!$, even when x is not a positive integer. The function $x!$ is also called the **factorial function**. The gamma function can also be defined as the solution of the functional equation $\Gamma(x+1) = x\Gamma(x)$ satisfying the conditions

$$\Gamma(1) = 1, \quad \lim_{n \rightarrow \infty} \frac{\Gamma(x+n)}{\Gamma(n)n^x} = 1.$$

Furthermore, we have

$$\frac{1}{\Gamma(x)} = x e^{Cx} \prod_{n=1}^{\infty} \left(1 + \frac{x}{n}\right) e^{-x/n}.$$

This expression is known as the **Weierstrass canonical form**. Here C is **Euler's constant**

$$C = \lim_{r \rightarrow \infty} \left(1 + \frac{1}{2} + \dots + \frac{1}{n} - \log n\right) \\ = 0.57721566490153286060651209 \dots,$$

which is conjectured to be [†]transcendental, but as yet even its irrationality remains unproved. However, it is known, that if C were rational, the numerator and the denominator would be integers of more than 30,000 digits (R. P. Brent, 1980).

The numerical value of C was calculated by Adams (1878) to 260 decimal places, and recently it has been calculated to more than 20,000 decimal places by means of an electronic computer. Seven thousand digits have been computed by W. A. Beyer and M. S. Waterman (*Math. Comp.*, 28 (1974)), and 20,000 digits have been computed by Brent et al. (1980).

$\Gamma(x)$ is [†]holomorphic on the complex x -plane except at the points $x = 0, -1, -2, \dots$, where it has simple [†]poles. When $\operatorname{Re} x > 0$, we have Hankel's integral representation

$$\Gamma(x) = -\frac{1}{2i \sin \pi x} \int_C (-t)^{x-1} e^{-t} dt,$$

$x \neq \text{integer},$

where the contour C lies in the complex plane cut along the positive real axis, starting at ∞ , going around the origin once counterclockwise, and ending at ∞ again.

Among various properties of this function (\rightarrow Appendix A, Table 17.I), the following two formulas are especially useful for numerical calculations: **Binet's formula**

$$\log \Gamma(x) = \left(x - \frac{1}{2}\right) \log x - x + \frac{\log 2\pi}{2} \\ + 2 \int_0^{\infty} \frac{\arctan(t/x)}{e^{2\pi t} - 1} dt, \quad \operatorname{Re} x > 0,$$

and the [†]asymptotic expansion formula that holds when $|\arg x| \leq (\pi/2) - \delta$ ($\delta > 0$),

$$\log \Gamma(x) \sim \left(x - \frac{1}{2}\right) \log x - x + \frac{\log 2\pi}{2} \\ + \sum_{n=1}^{\infty} \frac{(-1)^{n-1} B_{2n}}{2n(2n-1)x^{2n-1}}$$

(**Stirling's formula**), where the B_n are [†]Bernoulli numbers. This last formula can be rewritten as

$$\Gamma(x+1) = x! \sim x^x e^{-x} \sqrt{2\pi x},$$

which is used for large positive integers x .

The integrals

$$\int_0^{\lambda} e^{-t} t^{x-1} dt, \quad \int_{\lambda}^{\infty} e^{-t} t^{x-1} dt, \quad \operatorname{Re} x > 0,$$

are known as the **incomplete gamma functions** and are used in statistics, the theory of molecular structure, and other fields. The [†]exponential integral and the [†]error function (\rightarrow 167 Functions of Confluent Type D) are special cases of the incomplete gamma function.

B. Polygamma Functions

The derivatives of the logarithm of the gamma function are named the **digamma function** (or **psi function**) $\psi(x)=d\log\Gamma(x)/dx$; the **trigamma function** $\psi'(x)$; the **tetragamma function** $\psi''(x)$; the **pentagamma function** $\psi'''(x)$, etc. These functions are called **polygamma functions**. In particular, $\psi(x)$ is the solution of the functional equation

$$\psi(x+1)-\psi(x)=1/x,\quad \psi(1)=-C,$$
$$\lim_{n\rightarrow\infty}(\psi(x+n)-\psi(1+n))=0.$$

C. The Beta Function

Euler's integral of the first kind

$$B(x,y)=\int_0^1t^{x-1}(1-t)^{y-1}dt,$$

$\operatorname{Re} x>0,\quad \operatorname{Re} y>0,$

is called the **beta function** and is an analytic function of two variables x, y . This function is related to the gamma function as follows:

$$B(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

If the upper limit 1 in the integral is replaced by α , the result is called the **incomplete beta function** $B_\alpha(x,y)$.

References

[1] N. Nielsen, *Die Gammafunktion* I, II, Chelsea, 1964 (two volumes in one).
[2] W. Shibagaki, *Theory and applications of the gamma function, with tables of complex arguments* (in Japanese), Iwanami, 1952.
[3] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Cambridge Univ. Press, fourth edition, 1958.
[4] K. Pearson, *Tables of the incomplete Γ -function*, Cambridge Univ. Press, second edition, 1968.
[5] E. Artin, *The gamma function*, Holt, Rinehart and Winston, 1964. (Original in German, 1931.)
Also \rightarrow references to 389 Special Functions.

175 (XXI.26)
Gauss, Carl Friedrich

Carl Friedrich Gauss (April 30, 1777–February 23, 1855) was born into a poor family in Braunschweig, Germany. From

childhood, Gauss showed genius in mathematics. He gained the favor of Grand Duke Wilhelm Ferdinand and under his sponsorship attended the University of Göttingen. In 1797, on proving the †fundamental theorem of algebra, he received his doctorate from the University of Halle. From 1807 until his death, he was a professor and director of the Observatory at the University of Göttingen.

On March 30, 1796, he made the discovery that it is possible to draw a 17-sided †regular polygon with ruler and compass, which motivated his decision to devote himself to mathematics. The publication of his *Disquisitiones arithmeticae* in 1801 opened an entirely new era in number theory. In pure mathematics, he did excellent research on †non-Euclidean geometry, †hypergeometric series, the theory of functions of a complex variable, and the whole theory of †elliptic functions.

In the field of applied mathematics he made outstanding contributions to astronomy, geodesy, and electromagnetism; he also studied the †method of least squares, the theory of surfaces (\rightarrow 111 Differential Geometry of Curves and Surfaces), and the theory of †potential. He considered perfection in papers for publication to be of utmost importance; thus his published works are few relative to his amount of research. However, the scope of his work can be seen in his diary and letters, some of which are included in his complete works, comprising 12 volumes. He is generally considered the greatest mathematician of the first half of the 19th century.

References

[1] C. F. Gauss, *Werke* I–XII, Königliche Gesellschaft der Wissenschaften, Göttingen, 1863–1933.
[2] C. F. Gauss, *Disquisitiones arithmeticae*. German translation, *Untersuchungen über höhere Arithmetik*, Springer, 1889 (Chelsea, 1965); English translation, Yale Univ. Press, 1966.
[3] F. Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert* I, II, Springer, 1926–1927 (Chelsea, 1956).

176 (XVII.13)
Gaussian Processes

A. Gaussian Systems

A system $\mathbf{X}=\{X_\lambda(\omega)|\lambda\in\Lambda\}$ of real-valued random variables on a probability space

(Ω, \mathbf{B}, P) is said to be **Gaussian**, or \mathbf{X} is a **Gaussian system**, if any finite linear combination of elements X_λ of \mathbf{X} is a * Gaussian random variable.

Any subsystem of a Gaussian system is again a Gaussian system. In particular, the joint distribution of (X_1, X_2, \dots, X_n) for any finite subsystem $\{X_j | 1 \leq j \leq n\}$ of a Gaussian system \mathbf{X} is a multidimensional Gaussian distribution. This distribution is supported on the whole of \mathbf{R}^n or on a hyperplane (at most $(n-1)$ -dimensional). Let m_j be the expectation of X_j , and let $V = (V_{j,k})$ be the (positive definite) * covariance matrix of $\{X_j\}$ given by

$$V_{j,k} = E\{(X_j - m_j)(X_k - m_k)\}, \quad 1 \leq j, k \leq n.$$

Suppose that the rank of the matrix V is r . Then the distribution of (X_1, X_2, \dots, X_n) is concentrated on an r -dimensional hyperplane of \mathbf{R}^n :

$$m + V\mathbf{R}^n, \quad m = (m_1, m_2, \dots, m_n).$$

When V is nondegenerate, that is, when $r = n$, the distribution is supported by the whole of \mathbf{R}^n and has a density function of the form

$$(2\pi)^{-n/2} |V|^{-1/2} \exp \left[-\frac{1}{2} (x - m) V^{-1} (x - m)' \right],$$

where $x = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n$, $|V|$ and V^{-1} are the determinant and the inverse, respectively, of V , and where $(x - m)'$ denotes the (column) vector transpose to the (row) vector $(x - m)$. The above expression is the general form of the density of an n -dimensional Gaussian distribution. The characteristic function $\varphi(z)$ of this distribution is given by

$$\varphi(z) = \exp \left[i(m, z) - \frac{1}{2} (Vz, z) \right],$$

$$z = (z_1, z_2, \dots, z_n) \in \mathbf{R}^n,$$

where (\cdot, \cdot) denotes the inner product on \mathbf{R}^n . This distribution is denoted by $N(m, V)$. If, in particular, $m = 0$ and if V is the identity matrix, then it is called the n -dimensional standard Gaussian distribution.

For a general Gaussian system $\mathbf{X} = \{X_\lambda | \lambda \in \Lambda\}$ we are given the mean vector $m_\lambda = E(X_\lambda)$, $\lambda \in \Lambda$, and the covariance matrix $V_{\lambda, \mu} = E\{(X_\lambda - m_\lambda)(X_\mu - m_\mu)\}$, $\lambda, \mu \in \Lambda$, which is positive definite; for any $n \geq 1$, complex numbers $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbf{C}$, and $\lambda_1, \lambda_2, \dots, \lambda_n \in \Lambda$, we have

$$\sum_{j,k=1}^n \alpha_j \bar{\alpha}_k V_{\lambda_j, \lambda_k} \geq 0.$$

Conversely, given m_λ , $\lambda \in \Lambda$, and a positive definite $V = (V_{\lambda, \mu} | \lambda, \mu \in \Lambda)$, there exists a Gaussian system $\mathbf{X} = \{X_\lambda | \lambda \in \Lambda\}$, the mean vector and the covariance matrix of which coincide with (m_λ) and $V = (V_{\lambda, \mu})$, respectively. If there exists another system \mathbf{X}' with the same

property as \mathbf{X} , then \mathbf{X} and \mathbf{X}' have the same distribution.

A necessary and sufficient condition for the X_λ in a Gaussian system to be independent is that $V_{\lambda, \mu} = 0$ for every pair $\lambda \neq \mu$. For a Gaussian system $\mathbf{X} = \{X_n | n \geq 1\}$, the * convergence in probability of the sequence $\{X_n\}$ is equivalent to * convergence in the mean square. The limit of the sequence in this case is again Gaussian in distribution. Also for this system, the * almost sure convergence of the sequence implies the convergence in mean square.

There are many characterizations of Gaussian distributions and of Gaussian systems.

(i) A necessary and sufficient condition for a distribution to be Gaussian is that cumulants (* semi-invariants) γ_k of all orders exist and satisfy $\gamma_k = 0$ for all $k \geq 3$. (ii) Gaussian distributions have a self-reproducing property: If X and Y are independent Gaussian random variables, then their sum $S = X + Y$ is also Gaussian. A converse to this property holds: If the sum is Gaussian, then, assuming that they are independent, X and Y are both Gaussian as well (P. Lévy, H. Cramér). (iii) If X_1 and X_2 are independent and if $Y_1 = aX_1 + bX_1$ and $Y_2 = cX_1 + dX_2$ are also independent, then both X_1 and X_2 are Gaussian except for the trivial case $b = c = 0$ or $a = d = 0$. (iv) If for X and Y there exist U independent of X and V independent of Y satisfying $Y = aX + U$, $X = bY + V$ for some constants a, b , then there are only three possibilities: (1) (X, Y) is Gaussian, (2) X and Y are independent, (3) there is an affine relation between X and Y . (v) Suppose that a distribution has a finite mean m and a density function of the form $f(x - m)$. If the * maximum likelihood estimate of the mean is always given by the arithmetic mean of the samples, then the distribution is Gaussian (C. F. Gauss).

For any element X_λ of a Gaussian system \mathbf{X} and for a subsystem \mathbf{X}' of \mathbf{X} , the conditional expectation $E(X_\lambda | \mathbf{B}')$ is the orthogonal projection of X_λ onto $\overline{\mathbf{X}'}$, where \mathbf{B}' is the smallest σ -field with respect to which all the X_μ in \mathbf{X}' are measurable and where $\overline{\mathbf{X}'}$ is the closed linear subspace of $L^2(\Omega, P)$ spanned by \mathbf{X}' .

A system $\mathbf{X} = \{X_\lambda | \lambda \in \Lambda\}$, $X_\lambda = (X_\lambda^1, \dots, X_\lambda^d)$, of d -dimensional random variables is said to be Gaussian if the collection $\{X_\lambda^j | \lambda \in \Lambda, 1 \leq j \leq d\}$ is Gaussian.

B. Complex Gaussian Systems

Let Z be a complex-valued random variable with mean m , and denote it in the form $Z = X + iY + m$, $i = \sqrt{-1}$, X, Y real. If X and Y are independent and have the same Gaussian distribution with zero mean, then Z is called a

complex Gaussian random variable. A system $Z = \{Z_\lambda | \lambda \in \Lambda\}$ of complex-valued random variables is said to be **complex Gaussian** if any finite linear combination $\sum c_j Z_{\lambda_j}$, $c_j \in \mathbb{C}$, is complex Gaussian. A **complex Gaussian system** has properties similar to those of a Gaussian system discussed above. For instance, two complex Gaussian random variables are independent if and only if they are uncorrelated. Convergence properties are also similar. Furthermore, one can define complex Gaussian systems consisting of higher-dimensional random variables.

C. Gaussian Processes

A real-valued stochastic process $\{X_t\}$ is called a **Gaussian process** or a **normal process** if it forms a Gaussian system. If $\{X_t\}$ is a complex Gaussian system, it is called a complex Gaussian process. The most important example of a Gaussian process is [†]Brownian motion $\{B_t | t \geq 0\}$ with the properties $E(B_t) = 0$ and $E(B_t - B_s)^2 = |t - s|$. There are several generalizations of this process, among them: (i) [†]Brownian motion with a d -dimensional time parameter, (ii) a [†]Wiener process with d -dimensional time parameter, which is a Gaussian system $\{X_a | a = (a_1, \dots, a_d), \text{ all } a_j \geq 0\}$ with $E(X_a) = 0$ and $E(X_a, X_{a'}) = \prod_j \min\{a_j^1, a_j^2\}$, $a^i = (a_1^i, \dots, a_d^i)$, $i = 1, 2$.

Since a multidimensional Gaussian distribution is completely determined by its [†]mean vector and the [†]covariance matrix, a Gaussian process is strongly stationary if it is weakly stationary (\rightarrow 395 Stationary Processes); and such a process is called a **stationary Gaussian process**. The mean value m and the spectral measure $F(d\lambda)$ are associated with a weakly stationary process. The measure $F(d\lambda)$ is symmetric with respect to the origin since each X_t is real-valued. Conversely, given such $F(d\lambda)$ and m , we can construct a real weakly stationary process $\{X_t\}$ with mean value m and spectral measure $F(d\lambda)$. Generally, such a process $\{X_t\}$ is not determined uniquely; however, if $\{X_t\}$ is Gaussian, then there exists only one stationary Gaussian process with given m and $F(d\lambda)$. In view of this fact, stationary Gaussian processes can be regarded as being typical among weakly stationary processes.

The discussions so far on stationary Gaussian processes are generalized to the case of stationary complex Gaussian processes, for which symmetry of $F(d\lambda)$ need not be assumed. The [†]random measure $\{M(\Lambda)\}$ in the [†]spectral decomposition of a complex Gaussian process $\{X_t\}$ again forms a complex Gaussian system. Weakly stationary Gaussian random distributions can also be introduced.

D. White Noise and Gaussian Random Distributions

A real-valued weakly stationary process with discrete parameter is called a **white noise** if the mean $m = 0$ and the [†]covariance function $\rho(t) = 1$ ($t = 0$); $= 0$ ($t \neq 0$). Obviously the [†]spectral measure $F(d\lambda)$ is the [†]Lebesgue measure $d\lambda$. In the continuous parameter case, a weakly stationary random distribution (\rightarrow 407 Stochastic Processes C) is called a **white noise** if $m = 0$ and $\rho = \delta$ (δ is [†]Dirac's delta function: $\delta(\varphi) = \varphi(0)$). The spectral measure is, therefore, the Lebesgue measure. In both cases Gaussian white noise $\{X_t\}$ or $\{X_\varphi\}$ is determined uniquely. It has independent values at every point in the following sense: In the discrete parameter case, $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ are mutually independent for any mutually distinct t_1, t_2, \dots, t_n . For a continuous parameter Gaussian white noise, often called simply white noise if no confusion occurs, $X_{\varphi_1}, X_{\varphi_2}, \dots, X_{\varphi_n}$ are mutually independent if the supports of $\varphi_1, \varphi_2, \dots, \varphi_n$ are disjoint. In the latter case Gaussian white noise is realized by taking the derivative of a Brownian motion. The [†]characteristic functional $c(\varphi)$ is given by

$$c(\varphi) = \exp \left[-\frac{1}{2} \|\varphi\|^2 \right],$$

where $\| \cdot \|$ is the $L^2(\mathbb{R}^1)$ -norm.

Characteristic functionals of general Gaussian random distributions are expressed in the form

$$c(\varphi) = \exp \left[im(\varphi) - \frac{1}{2} K(\varphi, \varphi) \right],$$

where m is the mean functional and K is the covariance functional.

E. Representations of Gaussian Processes

The family of [†]stochastic integrals $X_t = \int_a^t F(t, u) dB_u$, $t > a$ ($t > a$ if $a = -\infty$), based on Brownian motion, defines a Gaussian process $\{X_t | t \geq a\}$. The converse problem is, however, not obvious. Given a Gaussian process $\{X_t | t > a\}$ with $E(X_t) = 0$, the problem is to find a Brownian motion $\{B_t\}$ and a kernel F that give a representation of the above form (P. Lévy [10]). A more specific problem discussed below is important. If, in addition, the representation is formed in such a way that $B_t(X) = B_t(B)$ holds for every t , then the representation is called **canonical**, where $B_t(X)$ is the smallest σ -field with respect to which all the X_s , $s \leq t$, are measurable. The **canonical representation**, if it exists, is unique up to equivalence, i.e., $F(t, u)^2$ is unique. The existence of the canonical representation is, however, not

always guaranteed. Mean continuous, zero mean, purely nondeterministic stationary Gaussian processes have †backward moving average representations, which are nothing but the canonical representations. Once the canonical representation of $\{X_t\}$ is obtained, the kernel, together with known properties of $\{B_t\}$, tells us important properties of the given process, e.g., sample function properties, Markov properties, and so forth.

Let $\{X_t | t > a\}$ be a general Gaussian process with $E(X_t) = 0$, and let $M_t(X)$ be the subspace of $L^2(\Omega, P)$ spanned by the X_s , $s \leq t$. Assume that (i) $\{X_t\}$ is separable, i.e., $M_\infty(X)$ is separable and (ii) $\{X_t\}$ is purely nondeterministic in the sense that $\bigcap_s M_s(X) = \{0\}$. Then X_t is expressed as a sum of stochastic integrals:

$$X_t = \sum_j \int_a^t F_j(t, u) dB_u^j,$$

where the $\{B_u^j\}$, $j = 1, 2, \dots, N$ (N can be ∞), are mutually independent †additive Gaussian processes. In addition, the B_u^j can be taken in such a way that the measures $dv^j(u) = E(dB_u^j)^2$, $j = 1, 2, \dots, N$ is decreasing in j (i.e., $dv^1 \gg dv^2 \gg \dots$) and that $B_t(X) = \bigvee_j B_t(B^j)$ holds for every t . Such a representation is called a **generalized canonical representation** of $\{X_t\}$ [3]. It always exists under the above assumptions (i) and (ii), although it is not unique when $N \geq 2$. The number N , called the **multiplicity** of $\{X_t\}$, is independent of the choice of a generalized canonical representation. It is a good question to ask when $\{X_t\}$ has a simple unit multiplicity or when $\{X_t\}$ has the canonical representation. No interesting answer to this question has been given so far except for stationary Gaussian processes.

F. Gaussian Markov Properties

A †simple Markov Gaussian process $\{X_t | t \geq a\}$ has the canonical representation if it is separable and purely nondeterministic and if its covariance function never vanishes. It is of the form

$$X_t = f(t) \int_a^t g(u) dB(u),$$

$$f(t) \neq 0, \quad \int_a^t g(u)^2 du \neq 0 \quad \text{for } t > a.$$

A generalization of the simple Markov property of a Gaussian process is given in the following manner (suggested by P. Lévy [10]). If the †conditional expectations $E(X_{t_k} | B_t(X))$, for any choice of distinct $t_1, t_2, \dots, t_{N+j} > t$ ($j \geq 0$) span an N -dimensional subspace of $L^2(\Omega, P)$, then the process $\{X_t\}$ is called an **N -ple Markov Gaussian process**. If the canonical representation of an N -ple Markov Gaussian

process exists, then it is expressed in the form

$$X_t = \sum_{j=1}^N f_j(t) \int_a^t g_j(u) dB_u,$$

where $\det(f_j(t_i))$ never vanishes for any choice of distinct t_i ($> a$), $1 \leq i \leq N$, and where the g_j are linearly independent vectors in $L^2([a, t])$ for any $t > a$. Conversely, a Gaussian process given by the above expression is N -ple Markov. A stationary N -ple Markov Gaussian process has a †spectral density function $f(\lambda)$ of a specific form, namely, it is expressed in the form [3]

$$f(\lambda) = |Q(i\lambda)/P(i\lambda)|^2; \quad P, Q \text{ polynomials,}$$

$$\deg P = N \quad \text{and} \quad \deg Q < N$$

(**rational spectral density function**). Y. Okabe [14] proved that the roots of $P(x) = 0$ are all real, and introduced a multiple Markov property of a stationary Gaussian process $\{X_t\}$ in a much wider sense to prove that $\{X_t\}$ enjoys this property if and only if it has a rational spectral density.

A somewhat restricted definition of multiple Markov properties for a Gaussian process uses differential operators. Assume that X_t is $(N-1)$ -times differentiable with respect to the $L^2(\Omega, P)$ -norm. If there is an N th order differential operator $L = \sum_{k=0}^N a_k(t) D^{N-k}$ with $D = d/dt$ such that

$$LX_t = \dot{B}_t, \quad \dot{B}_t = dB_t/dt \text{ white noise,}$$

then $\{X_t\}$ is called an **N -ple Markov Gaussian process in the restricted sense**. Such a process is naturally N -ple Markov, and the canonical representation always exists. The kernel of the representation is the †Riemann function of the differential operator L that was used in the definition. The spectral density function corresponding to this process is of restricted form, namely, it is expressed in the form $1/|P(i\lambda)|^2$, where P is a polynomial of degree N , and $P(i\lambda) = 0$ has no root in the lower half-plane.

Many attempts have been made to define a Markov property of a Gaussian random field, namely, a Gaussian system with a multi-dimensional time parameter; however, only two significant approaches are mentioned here. Let $\{X_a | a \in \mathbb{R}^d\}$ be a Gaussian random field with d -dimensional time parameter. H. P. McKean [12] gave a Markov property in the following manner. Let \mathcal{F}_U be the σ -field generated by the X_x , $x \in U$, $U (\subset \mathbb{R}^d)$ open. For a closed set C , we set $\mathcal{F}_C = \bigcap_\varepsilon \mathcal{F}_{C_\varepsilon}$, where C_ε is the ε -neighbourhood of C . If, for any open set U , \mathcal{F}_U and \mathcal{F}_{U^c} become independent under the assumption that $\mathcal{F}_{\partial U}$ is known, then $\{X_a\}$ is called a **Markov Gaussian random field in the McKean sense**. It has been proved that a Brownian motion with d -dimensional time

parameter is Markov in this sense if d is odd, while it is not Markov if d is even. Further detailed investigations of Markov properties for Gaussian random fields have been given by L. Pitt, S. Kotani, Y. Okabe [17], and others.

Another definition of Markov properties of Gaussian random fields, in fact those of Gaussian random distributions, has been given by E. Nelson [13] in connection with Euclidean fields (invariant under Euclidean motion) in quantum theory. Let X_Φ be a Gaussian random field. For a smooth surface Γ in \mathbf{R}^d let \mathcal{F}_Γ be the σ -field generated by X_Φ , $\Phi = \psi \otimes \delta(n)$, ψ being a smooth function with compact support and n being normal to Γ . If \mathcal{F}_D and \mathcal{F}_{D^c} become independent under the assumption that \mathcal{F}_Γ , $\Gamma = \partial D$, is given, then $\{X_\Phi\}$ is called a **Markov Gaussian random field in the Nelson sense**. Given a Gaussian Euclidean field which is Markov in the Nelson sense, then under some reasonable assumptions, one can form a field (in the sense of quantum field theory) satisfying the Wightman axiom by changing imaginary time to real time.

G. Sample Function Properties

Some finer results about the smoothness of sample functions have been obtained for stationary Gaussian processes. G. A. Hunt proved that if the spectral measure $F(d\lambda)$ of a stationary Gaussian process $\{X_t\}$ has finite moment of order α , then almost all its sample functions satisfy the following Hölder condition of order α [5]: For every finite interval I and every constant C , there exists a sufficiently small $h_0 = h_0(I, C)$ such that

$$|X_{t+h} - X_t| < C|h|^\alpha (\log 1/|h|)^{1/2}$$

for any $t \in I$ and $|h| < h_0$. The following result is due to Yu. K. Belyaev: Sample functions of a stationary Gaussian process are with probability 1 either continuous or unbounded on every finite interval.

Other conditions for continuity of sample functions are given in terms of the covariance function. Let $v(s, t)$ be the covariance function of a Gaussian process $\{X_t | t \in [0, 1]\}$. If v satisfies

$$|v(s_1, t) - v(s_2, t)| \leq c|s_1 - s_2|^\alpha$$

$$(0 < \alpha \leq 1, \quad v(0, s) = 0, \quad s_1, s_2 \in [0, 1]),$$

then

$$P\left(\limsup_{\delta \downarrow 0, |t_1 - t_2| \leq \delta} \frac{|X_{t_1} - X_{t_2}|}{|t_1 - t_2|^{1/2} |\log |t_1 - t_2||^{1/2}} < c_1 c^{1/2} / \alpha^2\right) = 1,$$

where c_1 is a constant. This result is due to Z. Ciesielski.

Conditions for continuity of sample functions have also been given for Gaussian processes with multidimensional time parameters. For $\{X_t | t \in [0, 1]^n\}$, set $d(s, t) = E[(X_s - X_t)^2]^{1/2}$. If there exists a function φ monotone increasing on some interval $0 < u < \alpha$ such that

$$d(s, t) \leq \varphi(\|s - t\|),$$

where $\| \cdot \|$ is the metric in \mathbf{R}^n , and such that

$$\int_0^{+\infty} \varphi(e^{-x^2}) dx < \infty,$$

then almost all sample functions are continuous (X. Fernique [2]). Conversely, if there exists a monotone increasing function φ satisfying $d(s, t) = \varphi(\|s - t\|)$ and if sample functions are continuous with positive probability, then the above integral should be finite.

Suppose that $\{X_t\}$ is a stationary Gaussian field. R. M. Dudley [11] proved that if there exists a number $q > 1$ and a neighborhood $V = V(0)$ in \mathbf{R}^n such that

$$\sum_k q^{-k} (\log N(V, q^{-k}))^{1/2} < \infty,$$

where $N(V, \varepsilon)$ is the minimum number of ε -balls (relative to the metric d) that cover V , then almost all sample fields of $\{X_t\}$ are continuous, and Fernique [18] proved that the converse statement is also true. Thus the problem of continuity of the sample fields of stationary Gaussian fields has been completely solved.

H. Gaussian Measures

Let $\{X_t, t \in T\}$, T an interval, be a Gaussian process. It gives a probability distribution on the measurable space $(\mathbf{R}^T, \mathcal{B})$, where \mathcal{B} is the σ -field of Borel subsets of \mathbf{R}^T . Such a measure P is called a Gaussian measure. Let P^i , $i = 1, 2$, be Gaussian measures induced by Gaussian processes $\{X_t^i, t \in T\}$, $i = 1, 2$, respectively. Then for these measures concepts such as absolute continuity, equivalence, and singularity are introduced in the usual manner. The most significant property for Gaussian measures is that two measures P^1 and P^2 are either equivalent or singular. A powerful criterion for testing this dichotomy is the entropy of P^2 with respect to P^1 , given by

$$H(P^2 | P^1) = \sup_\alpha \sum_i \left\{ P^2(A_i) \log \frac{P^2(A_i)}{P^1(A_i)} \right\},$$

where the sup is taken over all finite partitions $\alpha = \{A_1, \dots, A_n\}$, $A_i \in \mathcal{B}$, $\bigcup A_i = \mathbf{R}^T$ (Yu. A. Rozanov [15]). The measures P^1 and P^2 are equivalent (resp. singular) if and only if

$H(P^2|P^1) < \infty$ (resp. $= \infty$). This statement can be paraphrased in terms of the Hilbert spaces $H(X^i)$, $i = 1, 2$, spanned by $\{X_t^i\}$, $i = 1, 2$, respectively. For simplicity, assume that $E(X_t^i) = 0$. Then, P^1 and P^2 are equivalent if and only if the following two conditions hold: (i) A mapping A determined by $AX_t^1 = X_t^2$ is extended to an invertible linear transformation from $H(X^1)$ to $H(X^2)$; (ii) There exists a symmetric operator B of Hilbert-Schmidt type such that $(A\xi, A\eta)_2 = ((I - B)\xi, \eta)_1$, where $(\xi, \eta)_i = E(\xi\eta)$ is the inner product in $H(X^i)$. Next, suppose that $\{X_t^i\}$, $i = 1, 2$, are not centered. Set $E(X_t^i) = m_t^i$, $i = 1, 2$. Then P^1 and P^2 are equivalent if and only if $\{\bar{X}_t^i\}$ with $\bar{X}_t^i = X_t^i - m_t^i$ satisfy the following condition: (iii) in addition to conditions (i) and (ii) above, there exists an element ξ_0 in $H(X^1)$ such that $m_t^1 - m_t^2 = E[\xi_0 X_t^1]$, $t \in T$.

Given equivalent Gaussian measures P^1 and P^2 on $(\mathbf{R}^T, \mathcal{B})$, $T = [0, t_0]$, it is in general not easy to obtain the \dagger Radon-Nikodym derivative; however, if one of them, say P^1 , is the \dagger Wiener measure, then one can obtain $\varphi = dP^1/dP^2$ explicitly (M. Hitsuda [4]). Assume that P^2 is derived from $\{X_t\}$. By assumption, X_t has the \dagger canonical representation

$$X_t = B_t - \int_0^t \left\{ \int_0^s l(s, u) dB_u + a(s) \right\} ds, \quad t \in T,$$

where $\{B_t\}$ is the \dagger Brownian motion from which P^1 is assumed to be derived and where $l \in L^2(T^2)$, $a \in L^2(T)$. With this notation, we can write

$$\varphi = \frac{dP_1}{dP_2} = \exp \left[\int_0^{t_0} \left\{ \int_0^s l(s, u) dB_u + a(s) \right\} dB_s - \frac{1}{2} \int_0^{t_0} \left(\int_0^s l(s, u) dB_u + a(s) \right)^2 ds \right].$$

This result yields a remark about general Gaussian measures: If P^1 and P^2 are equivalent, $\{X_t^1, t \in T\}$ has a canonical representation if and only if $\{X_t^2, t \in T\}$ does.

1. Nonlinear Functionals

Start with nonlinear functionals of a Brownian motion $\{B_t, t \in \mathbf{R}^1\}$, and call them **Brownian functionals**. If $\{\varphi_n\}$ is a complete orthonormal system in $L^2(\mathbf{R}^1)$, then the collection $\{X_n = \int \varphi_n(u) dB_u\}$ of \dagger stochastic integrals forms a Gaussian system of independent standard Gaussian random variables. A **Fourier-Hermite polynomial** based on $\{\varphi_n\}$ is a polynomial in X_n expressible as a finite product in the form

$$c \prod_j H_{n_j}(X_j/\sqrt{2}), \quad c \text{ a constant.}$$

The sum $n = \sum n_j$ is the degree of this polynomial. If $n > 0$, its mean is zero and the variance

is $|c|^2 \prod_j (n_j! 2^{n_j})$. Taking c to be $\prod_j (n_j! 2^{n_j})^{-1/2}$, the collection of all Fourier-Hermite polynomials based on $\{\varphi_n\}$ forms a complete orthonormal system in the Hilbert space (L^2) consisting of all Brownian functionals with finite variance. Denote by \mathcal{H}_n the subspace of (L^2) spanned by all the Fourier-Hermite polynomials of degree n . Then the Hilbert space (L^2) admits a direct-sum decomposition:

$$(L^2) = \sum_{n=0}^{\infty} \oplus \mathcal{H}_n,$$

which is called the **Wiener-Itô decomposition** [7]. The subspace \mathcal{H}_n is eventually independent of the choice of a complete orthonormal system $\{\varphi_n\}$. A member X of \mathcal{H}_n is referred to as a **multiple Wiener integral** of degree n and is expressed in the form

$$X = \int_{\mathbf{R}^n} \dots \int f(s_1, s_2, \dots, s_n) dB_{s_1} dB_{s_2} \dots dB_{s_n},$$

where f is a symmetric $L^2(\mathbf{R}^n)$ -function. In addition, $E(X^2) = n! \|f\|^2$, $\|\cdot\|$ being the $L^2(\mathbf{R}^n)$ -norm.

Brownian functionals can be expressed as functionals of white noise, so that the Hilbert space $L^2(\mu)$ is viewed as a realization of (L^2) , where μ is the probability distribution of the Gaussian white noise introduced in Section D.

Nonlinear functionals of a general Gaussian process can be dealt with in a similar but somewhat more complicated manner. If the process has a canonical representation, then the functionals can easily be rewritten as Brownian functionals.

J. Applications to Prediction Theory

If a Gaussian process $\{X_t | t \geq a\}$ has the canonical representation

$$X_t = \int_a^t F(t, u) dB_u,$$

then the best \dagger predictor $E(X_t | \mathcal{B}_s(X))$, $s < t$, is given by

$$\int_a^s F(t, u) dB_u.$$

This is in agreement with the best \dagger linear predictor.

No systematic approach for nonlinear prediction theory has been established so far. We illustrate this theory only in a typical case that arises from a stationary Gaussian process. Let $\{X_t\}$ be a real-valued stationary Gaussian process. Without loss of generality, we can assume that $E(X_t) = 0$. We consider the continuous parameter case, since the discrete parameter case is easier and can be inferred from it.

Let $\mathbf{M}_t(X)$ be the subspace of $L^2(\Omega, P)$ de-

finned in the same way as in Section E, and let $L_t = L^2(\Omega, \mathbf{B}_t(X))$. If the process is purely non-deterministic in the sense that $\bigcap_t \mathbf{M}_t(X) = \{0\}$, then X_t has the backward moving average representation

$$X_t = \int_{-\infty}^t f(t-s) dB_s,$$

where $\{B_t\}$ is a Brownian motion. Furthermore, this representation is canonical, i.e., $\mathbf{B}_t(X) = \mathbf{B}_t(B)$. Since \mathbf{B} is assumed to be $\bigvee_t \mathbf{B}_t$, every Y in (L^2) can be expressed in the form

$$Y = \sum_{p=0}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{s_3} \int_{-\infty}^{s_2} f_p(s_1, s_2, \dots, s_p) dB_{s_1} dB_{s_2} \dots dB_{s_p},$$

in terms of multiple Wiener integrals. The predictor of Y based on the known values $X_s, s \leq 0$, that is, the projection of Y on L_0 is given by truncating the domain of the integral as

$$\sum_{p=0}^{\infty} \int_{-\infty}^0 \dots \int_{-\infty}^{s_3} \int_{-\infty}^{s_2} f_p(s_1, s_2, \dots, s_p) dB_{s_1} dB_{s_2} \dots dB_{s_p}.$$

K. Extrapolation and Interpolation

Extrapolation of a stationary Gaussian process is used to find the best estimate of the unknown values of a given process when we have been given the only some of the past values of the process. M. G. Kreĭn [8] obtained some results by using the theory of the inverse spectral problem discussed by I. M. Gelfand and B. Levitan.

In contrast to extrapolation is the interpolation of a stationary Gaussian process. There is an important contribution to this problem by H. Dym and H. P. McKean [1] who developed the Kreĭn theory by applying the de Branges method to the theory of entire functions.

References

- [1] H. Dym and H. P. McKean, Gaussian processes, function theory, and the inverse spectral problem, Academic Press, 1976.
- [2] X. Fernique, Continuité des processus gaussiens, C. R. Acad. Sci. Paris, 258 (1964), 6058–6060.
- [3] T. Hida, Canonical representations of Gaussian processes and their applications, Mem. Coll. Sci. Univ. of Kyoto, (A, Math.) 34 (1960), 109–155.
- [4] M. Hitsuda, Representation of Gaussian processes equivalent to Wiener processes, Osaka J. Math., 5 (1968), 299–312.

- [5] G. A. Hunt, Random Fourier transforms, Trans. Amer. Math. Soc., 71 (1951), 38–69.
- [6] I. A. Ibragimov and Yu. A. Rozanov, Gaussian random processes, Springer, 1978. (Original in Russian, 1970.)
- [7] K. Itô, Multiple Wiener integral, J. Math. Soc. Japan, 3 (1951), 157–169.
- [8] M. G. Kreĭn, On the fundamental approximation problem in the theory of extrapolation and the filtration of stationary random processes, Selected Transl. in Math. Stat. Prob., 4 (1963), 127–131. (Original in Russian, 1954.)
- [9] P. Lévy, Processus stochastiques et mouvement brownien, Gauthier-Villars, 1948; second edition, 1965.
- [10] P. Lévy, A special problem of Brownian motion and a general theory of Gaussian random functions, Proc. 3rd Berkeley Symp. Math. Stat. Prob. II, Univ. of Calif. Press, 1956, 133–175.
- [11] R. M. Dudley, The size of compact subset of Hilbert space and continuity of Gaussian processes, J. Functional Anal., 1 (1967), 290–330.
- [12] H. P. McKean, Jr., Brownian motion with a several-dimensional time, Teor. Veroyatnost i Primenen., 8 (1963), 357–378.
- [13] E. Nelson, Construction of quantum fields from Markov fields, J. Functional Anal., 12 (1973), 97–112.
- [14] Y. Okabe, On the structure of splitting fields of stationary Gaussian processes with finite multiple Markovian property, Nagoya Math. J., 54 (1974), 191–213.
- [15] Yu. A. Rozanov, Infinite-dimensional Gaussian distribution, Proc. Steklov Inst. Math., 108 (1971). (Original in Russian, 1968.)
- [16] M. Yor, Existence et unicité de diffusion à valeurs dans un espace Hilbert, Ann. Inst. H. Poincaré, (B) 10 (1974), 55–88.
- [17] Y. Okabe, Stationary Gaussian processes with Markovian property and M. Sato's hyperfunctions, Japan J. Math., 41 (1973), 69–122.
- [18] X. Fernique, Régularité des trajectoires des fonctions aléatoires gaussiennes, Lecture notes in math. 480, Springer, 1974, 1–96.

177 (XIV.2) Generating Functions

A. General Remarks

A power series $g(t) = \sum_{n=0}^{\infty} a_n t^n$ in t that converges in a certain neighborhood of $t=0$ defines the sequence of numbers $\{a_n\}$. The function $g(t)$ is called the **generating function** of

the sequence. Similarly, the series $K(x, t) = \sum_{n=0}^{\infty} f_n(x) t^n$ that is convergent for x and t in a certain domain in (x, t) -space is called the **generating function** of the sequence of functions $\{f_n(x)\}$. Sometimes the function $g(t) = \sum_{n=0}^{\infty} (a_n/n!) t^n$ is called the **exponential generating function** of the sequence $\{a_n\}$. For example, the generating functions of the \dagger binomial coefficients and \dagger Legendre polynomials are $(1+t)^n$ and $(1-2tx+t^2)^{-1/2}$, respectively. When a generating function of $\{a_n\}$ or $\{f_n(x)\}$ is given, we can obtain a_n or $f_n(x)$ by integral expressions; for example, in the latter case we have

$$f_n(x) = \frac{1}{2\pi i} \int_C \frac{K(x, t)}{t^{n+1}} dt,$$

where the contour C is a sufficiently small circle going counterclockwise around the origin. A generating function can be continued beyond the domain of convergence of the power series. Simple generating functions are known for many important orthogonal systems of functions (\rightarrow 317 Orthogonal Functions). Generating functions are widely used because they enable us to derive analytically the properties of sequences of numbers or functions. For a system of numbers or functions depending on a continuous parameter instead of the integral parameter n , we define the generating function in the form of a \dagger Laplace or \dagger Fourier transform. In particular, for a probability \dagger distribution function $F(x)$, the exponential generating function $f(t) = \int_{-\infty}^{\infty} e^{-tx} dF(x)$ for the moments $\{a_n\}$ of $F(x)$ is called the **moment-generating function** of $F(x)$. There is another kind of generating function for the sequence $a_n: \sum a_n n^{-s}$ in the form of \dagger Dirichlet series; it is frequently used in number-theoretic problems.

B. Bernoulli Polynomials

A system of polynomials

$$B_n(x) = \sum_{k=0}^n \binom{n}{k} B_k(0) x^{n-k}$$

is defined by the generating function

$$\frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{t^n}{n!}.$$

$B_n(x)$ is called the **Bernoulli polynomial** of degree n . Since $B_k(0)$ is the coefficient of $t^k/k!$ in

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n(0) \frac{t^n}{n!},$$

we have $B_0(0) = 1$, $B_1(0) = -1/2$, $B_2(0) = 1/6$, ...; $B_{2n+1}(0) = 0$ for $n \geq 1$; and $(-1)^{n-1} B_{2n}(0) > 0$ for $n \geq 1$. The n th **Bernoulli number** B_n (\rightarrow Ap-

pendix B, Table 3) is usually defined by $|B_n(0)|$. Sometimes other definitions, such as $B_n = B_n(0)$ or $B_n = B_{2n}(0)$, are used. Bernoulli polynomials satisfy the relations

$$B_n(x+1) - B_n(x) = nx^{n-1},$$

$$dB_n(x)/dx = nB_{n-1}(x),$$

which are used in \dagger interpolation. For example, a polynomial solution of the \dagger difference equation $f(x+1) - f(x) = \sum_{n=0}^m a_n x^n$ is given by $f(x) = \sum_{n=0}^m a_n B_{n+1}(x)/(n+1) + (\text{arbitrary constant})$. In particular, we have $1^n + 2^n + \dots + p^n = (B_{n+1}(p+1) - B_{n+1}(1))/(n+1)$.

C. Euler Polynomials

A system of polynomials

$$E_n(x) = \sum_{k=0}^n \binom{n}{k} a_k x^{n-k}$$

is defined by the generating function

$$\frac{2e^{xt}}{e^t + 1} = \sum_{n=0}^{\infty} E_n(x) \frac{t^n}{n!}.$$

We call $E_n(x)$ the **Euler polynomial** of degree n . Here a_k is defined by $a_k = E_k(0)$ and

$$\frac{2}{e^t + 1} = \sum_{n=0}^{\infty} a_n \frac{t^n}{n!},$$

so that we have $a_0 = 1$, $a_1 = -1/2$, $a_3 = 1/4$, ...; $a_{2n} = 0$ for $n \geq 1$. The n th Euler number is sometimes taken as a_n , but more often it is defined by

$$E_n = (-1)^n \sum_{\mu=0}^n 2^\mu \binom{n}{\mu} a_\mu,$$

i.e., by

$$\frac{2}{e^x + e^{-x}} = \text{sech } x = \sum_{n=0}^{\infty} (-1)^n \frac{E_n}{n!} x^n$$

(\rightarrow Appendix B, Table 3). All the E_n are integers, $E_{2m+1} = 0$ ($m = 0, 1, \dots$) and $E_{2m} > 0$ ($m = 0, 1, \dots$); in the decimal expressions for E_n the last digit is 5 for E_{4m} ($m \geq 1$) and 1 for E_{4m+2} ($m \geq 0$). Sometimes E_{2n} is denoted by E_n . We have the relations

$$\sec x = \sum_{n=0}^{\infty} \frac{E_n x^n}{n!},$$

$$E_n(x) + E_n(x+1) = 2x^n,$$

$$E_n(1-x) = (-1)^n E_n(x),$$

$$\frac{dE_n(x)}{dx} = nE_{n-1}(x),$$

and in particular,

$$\begin{aligned} -1^n + 2^n - 3^n + 4^n - \dots + (-1)^p p^n \\ = ((-1)^p E_n(p+1) - E_n(1))/2. \end{aligned}$$

D. Application to Combinatorics

Let $p(n)$ denote the number of ways of dividing n similar objects into nonempty class. This is called the **number of partitions** of n . Euler noticed that the following formula is valid for the generating function of $p(n)$:

$$1 + \sum_{n=1}^{\infty} p(n)x^n = \left[\prod_{n=1}^{\infty} (1 - x^n) \right]^{-1},$$

whence we obtain $p(n) = p(n-1) + p(n-2) - p(n-5) - p(n-7) + p(n-12) + p(n-15) - \dots = \sum_{k=1}^{\infty} (-1)^{k-1} [p(n - (3k^2 - k)/2) + p(n - (3k^2 + k)/2)]$ (\rightarrow 328 Partitions of Numbers). Let $B(n)$ denote the number of ways of dividing n completely dissimilar objects into nonempty classes. These are called **Bell numbers**. The first few are 1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147, ... [2]. For this sequence with $B(0) = 1$, the formal generating function $\sum_{n=0}^{\infty} B(n)x^n$ does not converge except at $x=0$. However, the exponential generating function $g(x) = \sum_{n=0}^{\infty} B(n)(n!)^{-1}x^n$ is convergent for all complex numbers x and is equal to e^{e^x-1} . The differential equation $g' = e^x g$ gives rise to a recursive formula

$$B(n+1) = \sum_{k=0}^n \binom{n}{k} B(k).$$

References

- [1] E. Netto, The theory of substitutions and its applications to algebra, revised and translated by F. N. Cole, second edition, Chelsea, 1964. (Original in German, 1882.)
 - [2] E. T. Bell, Algebraic arithmetic, Amer. Math. Soc., 1927.
 - [3] E. B. McBride, Obtaining generating functions, Springer, 1971.
- Also \rightarrow references to 389 Special Functions.

178 (VII.6) Geodesics

A. General Remarks

In the study of global differential geometry, geodesics play an essential role because the behavior of geodesics on a complete Riemannian m -manifold ($m \geq 2$) M without boundary heavily influences its topological structure. A significant merit of using geodesics is that one can make elementary and intuitive observations on M (as in Euclidean geometry), which often yields fruitful results.

Riemannian structure induces in a natural way a distance function d on M , and hence M is a metric space. A smooth curve γ of constant

speed is by definition a **geodesic** if and only if every point on γ is contained in a nontrivial subarc whose length realizes the distance between its endpoints. M is equipped with the Levi-Civita connection ∇ (\rightarrow 80 Connections) by means of which a geodesic is characterized as an autoparallel curve. In local coordinates, γ is obtained as a solution of an ordinary differential equation of order two (\rightarrow 80 Connections L). It follows from the properties of solutions of the equation for geodesics that every point $p \in M$ has a neighborhood in which every two points can be joined by a unique shortest geodesic that depends smoothly on both of the endpoints. A classical result of J. H. C. Whitehead (*Quart. J. Math.*, 3 (1932)) states that every point $p \in M$ has a metric ball B centered at p such that every two points in B are joined by a unique shortest geodesic whose image is contained in B . Let M_p be the tangent space to M at p , and let $\tilde{M}_p \subset M_p$ be the set of all vectors $v \in M_p$ such that the solution of geodesic γ with initial condition $\gamma(0) = p$, $\dot{\gamma}(0) = v$ is well defined on $[0, 1]$. \tilde{M}_p contains an open neighborhood of the origin and is star-shaped with respect to the origin. The exponential mapping $\exp_p: \tilde{M}_p \rightarrow M$ is defined to be $\exp_p v = \gamma(1)$. This mapping \exp_p is smooth and has the maximal rank at the origin. Small balls around p are obtained as the image under the exponential mapping of the corresponding balls in \tilde{M}_p centered at the origin, and the restriction of \exp_p to these balls are diffeomorphisms. Thus the topology of M as a metric space is equivalent to the original one of M . The fundamental Hopf-Rinow theorem (*Comment. Math. Helv.*, 3 (1931)) states that (1) M is complete as a metric space if and only if $\tilde{M}_p = M_p$ for some $p \in M$; (2) M is complete if and only if every closed metric ball is compact; and (3) if M is complete, then every two points can be joined by a shortest geodesic, namely, a geodesic with length realizing the distance between the endpoints.

In the following discussion M is assumed to be connected, complete, and without boundary. Let $\gamma: [0, 1] \rightarrow M$ be a geodesic. A piecewise smooth 1-parameter variation V along γ is a continuous mapping $V: [0, 1] \times (-\epsilon, \epsilon) \rightarrow M$ with a finite partition $0 = t_0 < t_1 < \dots < t_k = 1$ such that $V|_{[t_i, t_{i+1}] \times (-\epsilon, \epsilon)}$ is smooth and $V(t, 0) = \gamma(t)$ for all $t \in [0, 1]$. Then the first and second variation formulas for V are

$$L'(0) = \langle Y, \dot{\gamma} \rangle|_0^1 / L(\gamma)$$

and

$$L''(0) = \sum_{i=1}^k \left[\int_{t_{i-1}}^{t_i} (\langle Y_i', Y_i' \rangle - \langle R(Y_i, \dot{\gamma}) \dot{\gamma}, Y_i \rangle) dt + \langle \nabla_{\partial/\partial s} Y_i, \dot{\gamma} \rangle|_{t_{i-1}}^{t_i} \right] / L(\gamma),$$

respectively, where $Y_i(t) = \partial V(t, s)/\partial s|_{s=0}$ for $t \in [t_{i-1}, t_i]$ is the variation vector field associated with $V|_{[t_{i-1}, t_i] \times (-\varepsilon, \varepsilon)}$, $L(s)$ is the length of the variation curve $t \rightarrow V(t, s)$, and $Y'_i = \nabla_{\dot{\gamma}} Y_i$ and R is the curvature tensor.

A smooth vector field J along a geodesic $\gamma: [0, 1] \rightarrow M$ is called a **Jacobi field** if and only if it satisfies $J'' + R(J, \dot{\gamma})\dot{\gamma} = 0$. The set of all Jacobi fields along γ forms a vector space isomorphic to $M_{\gamma(0)} \times M_{\gamma(1)}$ by the natural correspondence $J \rightarrow (J(0), J'(0))$. Every Jacobi field is associated with a 1-parameter geodesic variation V along γ . Namely, every variation curve of V is a geodesic, V is smooth, and $J(t) = \partial V(t, s)/\partial s|_{s=0}$ for all $t \in [0, 1]$. Conversely, the variation vector field of any 1-parameter geodesic variation V along γ is a Jacobi field. Especially, if A is a tangent vector to M_p at $\dot{\gamma}(0)$, then $d(\exp_p)_{\dot{\gamma}(0)} A = J(1)$, where J is the Jacobi field associated with the 1-parameter geodesic variation $V(t, s) = \exp_p(t(\dot{\gamma}(0) + sA))$, and A is identified under the canonical parallel translation in M_p . A point $\gamma(t_0)$ is said to be **conjugate** to $p = \gamma(0)$ along γ if and only if there exists a nontrivial Jacobi field along γ that vanishes at 0 and t_0 . A theorem of Morse and Schoenberg states that if the sectional curvature K of M satisfies $0 < A \leq K \leq B < \infty$, then every unit speed geodesic $\gamma: [0, l] \rightarrow M$ has a point conjugate to $\gamma(0)$ if $l \geq \pi/\sqrt{A}$ and has no point conjugate to $\gamma(0)$ if $l < \pi/\sqrt{B}$. Especially, every geodesic γ on M with $K \leq 0$ contains no conjugate point on it, and \exp_p is locally regular for any $p \in M$. Thus a complete and simply connected M with $K \leq 0$ is diffeomorphic to \mathbf{R}^m (Hadamard and Cartan).

Let γ be a unit speed geodesic with $\gamma(0) = p$. If $t_1 > 0$ satisfies $d(p, \gamma(t)) = t$ for all $t \in [0, t_1]$ and $d(p, \gamma(t)) < t$ for all $t > t_1$, then the point $\gamma(t_1)$ is called a **cut point** to p along γ . It appears no later than the first conjugate point to p along γ . The set $C(p)$ of all cut points to p is called the **cut locus** of p . Let U be the set of vectors $tv \in M_p$, $0 \leq t < 1$, where $\exp_p v$ is the cut point to p along $\gamma(t) = \exp_p tv$. U is a nonempty open set and $\exp_p|_U: U \rightarrow M - C(p)$ is a diffeomorphism, where $M - C(p)$ is diffeomorphic to an m -disk. The cut locus possesses essential information on the topology of M .

The **metric comparison theorems** of H. Rauch (*Ann. Math.*, (2) 54 (1951)) and M. Berger (*Illinois J. Math.*, 6 (1962)) are stated as follows: Let $\inf_M K = \delta > -\infty$ and $\sup_M K = \Delta < \infty$, and let $\gamma: [0, 1] \rightarrow M$ be a nontrivial geodesic. Let $M(c)$ be a complete and simply connected space form of constant curvature c . Fix geodesics $\gamma_\delta, \gamma_\Delta: [0, 1] \rightarrow M(\delta), M(\Delta)$, respectively, with the same speed as γ . If J, J_δ , and J_Δ are Jacobi fields along γ, γ_δ , and γ_Δ respectively such that $J(0) = J_\delta(0) = J_\Delta(0) = 0$ and $\|J'(0)\| = \|J'_\delta(0)\| = \|J'_\Delta(0)\|$, then $\|J_\Delta(t)\| \leq$

$\|J(t)\| \leq \|J_\delta(t)\|$ holds for all $t \in [0, t_\Delta]$, where $\gamma_\Delta(t_\Delta)$ is the first conjugate point to $\gamma_\Delta(0)$ along γ_Δ (Rauch). If the Jacobi fields satisfy $\|J(0)\| = \|J_\delta(0)\| = \|J_\Delta(0)\|$ and $J'(0) = J'_\delta(0) = J'_\Delta(0) = 0$, then $\|J_\Delta(t)\| \leq \|J(t)\| \leq \|J_\delta(t)\|$ holds for all $t \in [0, t'_\Delta]$, where t'_Δ is the first zero point of J_Δ (Berger). They are often applied to compare curve lengths as follows. Let $c: I \rightarrow \mathbf{R}^m$ be a piecewise smooth curve with $\|c(t)\| \leq t_\Delta$. Fix points $p \in M, p_\delta \in M(\delta)$, and $p_\Delta \in M(\Delta)$ with fixed isometric identifications of the tangent spaces with \mathbf{R}^m . Then $L(\exp_{p_\Delta} \circ c) \leq L(\exp_p \circ c) \leq L(\exp_{p_\delta} \circ c)$. If P, P_δ , and P_Δ are unit parallel fields along corresponding geodesics $\sigma, \sigma_\delta, \sigma_\Delta: I \rightarrow M, M(\delta), M(\Delta)$, respectively, which have the same speed, and if $\langle P, \dot{\sigma} \rangle = \langle P_\delta, \dot{\sigma}_\delta \rangle = \langle P_\Delta, \dot{\sigma}_\Delta \rangle$, then for every piecewise smooth function $f: I \rightarrow [0, t'_\Delta]$ the curves $u(s) = \exp_{\sigma(s)} f(s)P(s)$ and $u_\delta(s), u_\Delta(s)$ defined in the same way on $M(\delta), M(\Delta)$, respectively, have lengths $L(u_\Delta) \leq L(u) \leq L(u_\delta)$. They play important roles in the theory of geodesics.

A geodesic triangle is a triple $(\gamma_0, \gamma_1, \gamma_2)$ of shortest geodesics (for convenience they are parametrized on $[0, 1]$) such that $\gamma_i(1) = \gamma_{i+1}(0)$ for all $i = 0, 1, 2$ with mod 3. The vertices p_0, p_1, p_2 and the angles $\theta_0, \theta_1, \theta_2$ of a geodesic triangle $(\gamma_0, \gamma_1, \gamma_2)$ are defined by $p_{i+1} = \gamma_i(0)$ and $\theta_{i+1} = \cos^{-1}\{\langle \dot{\gamma}_i(0), -\dot{\gamma}_{i-1}(1) \rangle / \|\dot{\gamma}_i\| \|\dot{\gamma}_{i-1}\|\}$, where angles are always taken in $[0, \pi]$. The **triangle comparison theorem** of Toponogov (*Amer. Math. Soc. Transl.*, 37 (1964)) is stated as follows. Let $K \geq \delta > -\infty$ be satisfied on M . For any geodesic triangle $(\gamma_0, \gamma_1, \gamma_2)$ on M there is a geodesic triangle $(\tilde{\gamma}_0, \tilde{\gamma}_1, \tilde{\gamma}_2)$ on $M(\delta)$ such that $L(\tilde{\gamma}_i) = L(\gamma_i)$ and the angles $\tilde{\theta}_i$ of this triangle satisfy $\tilde{\theta}_i \leq \theta_i$ for $i = 0, 1, 2$. It turns out that if $\delta > 0$ then the circumference of any geodesic triangle on M does not exceed $2\pi/\sqrt{\delta}$.

For details of the basic facts stated above we refer the reader to [1–3].

B. Curvature and Fundamental Groups

The fundamental group $\pi_1(M)$ of M (\rightarrow 170 Fundamental Groups) is influenced by the curvature of M . A basic idea, going back to Hadamard (*J. Math. Pures Appl.*, 4 (1898)) and Cartan, states that every nontrivial free homotopy class $\delta \in \pi_1(M)$ of loops on a compact M contains a closed geodesic of minimal length whose preimage under the covering projection $\pi: \tilde{M} \rightarrow M$ in the universal Riemannian covering \tilde{M} is either closed (when δ is of finite order) or a straight line (when δ is of infinite order) that is translated along itself by the deck transformation δ .

(1) By using the second variational formula, an even-dimensional compact M with $K > 0$

either is simply connected or has $\pi_1(M) = \mathbf{Z}_2$ (Synge, *Quart. J. Math.*, 7 (1936)). A beautiful result of Myers (*Duke Math. J.*, 8 (1941)) states that if the Ricci curvature Ric of M is bounded below by $\delta > 0$, then the diameter $d(M)$ is not greater than $\pi/\sqrt{\delta}$ and hence M is compact and $\pi_1(M)$ is finite. It is proved in [3] that if $K \geq 0$, then there exists a finite normal subgroup Φ of $\pi_1(M)$ such that $\pi_1(M)/\Phi$ is a Bieberbach group (\rightarrow 92 Crystallographic Groups). The splitting theorem (\rightarrow Section F(2)) is used in the proof.

(2) When $K \leq 0$ is satisfied on M , \tilde{M} is diffeomorphic to \mathbf{R}^m , and hence $\pi_i(M) = 0$ for all $i \geq 2$. A basic tool used in the study of $\pi_1(M)$ in this case is that the displacement function $\tilde{x} \rightarrow \tilde{d}(\tilde{x}, \delta(\tilde{x}))$, $\tilde{x} \in \tilde{M}$, of every isometry δ on \tilde{M} is convex [4], where a function on M is said to be **convex** if and only if its restriction to every geodesic is convex. A classical result of Preissmann (*Comment. Math. Helv.*, 15 (1943)) states that every nontrivial Abelian subgroup of $\pi_1(M)$ of a compact M with $K < 0$ is an infinite cyclic subgroup. It is proved in [4] that if $K < 0$ holds on M and if every deck transformation on \tilde{M} translates some geodesic along itself, then $\pi_1(M)$ is a disjoint union of infinite cyclic subgroups, and any two commuting elements belong to the same cyclic subgroup. Moreover, if M is compact and $K \leq 0$ and if $\pi_1(M)$ is a direct product $\Gamma_1 \times \Gamma_2$ such that $\pi_1(M)$ is \dagger centerless, then M is isometric to the Riemannian product $M_1 \times M_2$, and $\pi_1(M_i) = \Gamma_i$ holds for $i = 1, 2$ [5, 6]. It is shown in [7] that the fundamental group of M with $K \leq 0$ occurs as that of an $(m+1)$ -dimensional M' with $K \leq c < 0$, which is diffeomorphic to $M \times \mathbf{R}$. The warped product [4] is used to construct such a metric on $M \times \mathbf{R}$.

C. Cut Locus

The **injectivity radius** $i(M)$ of M is defined to be the infimum of the continuous function $x \rightarrow d(x, C(x))$, $x \in M$. As is seen in Section D, the estimate of injectivity radii provides many fruitful results on the topology of Riemannian manifolds. It follows from Synge's result that an even dimensional compact and orientable M with $0 < K \leq 1$ has its injectivity radius $i(M) \geq \pi$ [3]. However, such an estimate cannot be obtained in odd dimensions. The examples discussed in [8] show that there are infinitely many homotopically distinct homogeneous Riemannian 7-manifolds of positive curvature. The injectivity radii of such examples are estimated precisely in [9], according to which there is no positive lower bound for them. The **sphere theorem** of Klingenberg (*Comment. Math. Helv.*, 35 (1961))

states that if M is compact and simply connected and if $1/4 < K \leq 1$, then $i(M) \geq \pi$, and M is a topological sphere. This extends the pioneering work by H. Rauch (*Ann. Math.*, (2) 54 (1951)). A **rigidity theorem** of M. Berger (*Ann. Scuola Norm. Sup. Pisa*, 14 (1962)) states that if $\dim M$ is even, M is simply connected, and $1/4 \leq K \leq 1$, then M is homeomorphic to S^m (if $d(M) > \pi$) or isometric to one of the symmetric spaces of compact type of rank 1 (if $d(M) = \pi$). A slight generalization of the triangle comparison theorem is used to prove that for given positive constants d , V , and S , there exists a constant $C_m(d, V, S) > 0$ such that if K , $d(M)$ and the volume $v(M)$ satisfy $|K| < S$, $d(M) < d$, and $v(M) > V$, then $i(M) \geq C_m(d, V, S)$ [10].

D. Finiteness Theorem

Finiteness theorems are a natural extension of the sphere theorem. They provide *a priori* estimates for the number of various topological types of manifolds (for instance, homology, homotopy, homeomorphism, diffeomorphism types, etc.), which admit certain classes of Riemannian metrics characterized by geometric quantities. The basic idea of the estimates is to find a constant $c > 0$ that depends only on geometric information which characterizes the class of manifolds so that if M belongs to the class, then $i(M) \geq c$. Then a number N is found from the information given *a priori* such that every element M of the class has an open cover of at most N balls whose radii are all less than c . It is proved in [11] that for given $\delta \in (0, 1)$ and m , there exists a number $N(\delta, m)$ such that there are at most $N(\delta, m)$ homotopy types for the class of all simply connected $2m$ -dimensional manifolds with $\delta \leq K \leq 1$. Furthermore, for given positive numbers d , V , and S , there are at most finitely many homeomorphism (or diffeomorphism) types for the class of all m -manifolds M , each of which has the property that $d(M) < d$, $v(M) > V$, and $|K| < S$ [10]. This result is applied to obtain a result in [12], which states that there exists for given m , $V > 0$ and an $\varepsilon > 0$ such that if M is compact, $-1 - \varepsilon \leq K \leq -1$, and $v(M) > V$, then M admits a metric of constant negative curvature.

For an $\varepsilon \geq 0$, M is said to be ε -**flat** if and only if $\sup_M K \cdot d(M)^2 \leq \varepsilon$ is satisfied for M . Every compact flat manifold is ε -flat for all $\varepsilon \geq 0$. M is called a **nilmanifold** if and only if it admits a transitive action of a \dagger nilpotent Lie group. It is shown in [13] that for a compact M with $\dim M = m$, there exists a number $\varepsilon(m) > 0$ such that if M is $\varepsilon(m)$ -flat, then (1) there is a maximal nilpotent normal divisor $N \subset \pi_1(M)$, (2)

the order of $\pi_1(M)/N$ is bounded by a constant which depends only on m , and (3) the finite cover of M which corresponds to N is diffeomorphic to a nilmanifold. Moreover, if M is $\varepsilon(m)$ -flat then M is diffeomorphic to \mathbf{R}^m . If M is $\varepsilon(m)$ -flat and if $\pi_1(M)$ is commutative, then M is diffeomorphic to a torus.

E. Uniqueness Theorem

Uniqueness of topological structures (as in the sphere theorem) of certain classes of compact Riemannian manifolds is discussed here.

The discovery of exotic 7-spheres by J. Milnor (*Ann. Math.*, (2) 64 (1956)) gave rise to the question of whether in the sphere theorem the conclusion (homeomorphism to the sphere) could be replaced by diffeomorphism. Since the number of differentiable structures on a topological sphere depends on its dimension (\rightarrow 114 Differential Topology), it has been thought that in order to get the standard sphere in the sphere theorem the best possible restriction for the curvature might also depend upon dimension. The appropriate **differentiable pinching problem** is to find a sequence $\{\Delta_m\}$ so that if M is compact and simply connected and if $\delta \leq K \leq 1$ for some $\delta > \Delta_m$, then M is diffeomorphic to S^m , and Δ_m is the least possible with this property. D. Gromoll and Y. Shikata proved independently [14, 15] that $\Delta_m < 1$ holds for all $m \geq 2$. Later it was proved that there exists a $\delta_0 \in (1/4, 1)$ such that $\Delta_m \leq \delta_0$ holds for all $m \geq 2$ [16]. The **diffeotopy theorem**, which plays an essential role in the proof of finding such a δ_0 , provides a sufficient condition for a diffeomorphism on S^{m-1} to be isotopic to the identity mapping.

A different idea is put forth in [17], which imitates the Gauss normal mapping of a closed convex hypersurface in \mathbf{R}^{m+1} . It is proved that if the curvature operator is sufficiently closed to the identity, then M is diffeomorphic to S^m . This idea is used to obtain a diffeomorphism between M and a spherical space form. A generalization of the sphere theorem states that if M is compact and if $(d(M)/\pi)^2 \inf_M K > 1/4$, then M is a topological sphere [18].

F. Noncompact Manifolds

Let M be noncompact. It is due to the nature of noncompactness that through each point on M there passes a ray $\gamma: [0, \infty) \rightarrow M$, e.g., γ is a unit speed geodesic any of whose subarc is minimizing.

(1) Busemann Functions. A Busemann function $F_\gamma: M \rightarrow \mathbf{R}$ with respect to a ray γ is defined by

$F_\gamma(x) = \lim_{t \rightarrow \infty} [t - d(x, \gamma(t))]$. The original definition of it was used by H. Busemann to define a parallel axiom on straight G -spaces (\rightarrow Section H). F_γ has the property $F_\gamma^{-1}((-\infty, t_1]) = \{x \in F_\gamma^{-1}((-\infty, t_2]) \mid d(x, \partial F_\gamma^{-1}((-\infty, t_2])) \geq t_2 - t_1\}$ for any $t_2 \geq t_1$, and hence $F_\gamma(x) = t_2 - d(x, \partial F_\gamma^{-1}((-\infty, t_2]))$ for all t_2 and x with $F_\gamma(x) \leq t_2$. It has been proved, by using the second variation formula, that if $\text{Ric} \geq 0$, $K \geq 0$ [19] or if, in the case where M is Kählerian, the holomorphic bisectional curvature [20] is nonnegative, then F_γ is subharmonic (sh), convex or plurisubharmonic (psh), respectively, where the holomorphic bisectional curvature is defined as $R(X, JX, JY, Y)$ for the complex structure J and orthonormal vectors X and Y . If the holomorphic bisectional curvature [20] is positive, then F_γ is strictly psh. If $K \geq 0$, then the triangle comparison theorem implies that $F = \sup\{F_\gamma \mid \gamma(0) = p\}$ is convex and is an exhaustion [21], where the sup is taken over all rays emanating from a point $p \in M$, and a function f is said to be an **exhaustion** if and only if $f^{-1}((-\infty, a])$ is compact for all $a \in \mathbf{R}$. A well-known theorem of H. Grauert (*Math. Ann.*, 140 (1960)) states that if M admits a strictly psh exhaustion function, then M is a Stein manifold (\rightarrow 21 Analytic Functions of Several Complex Variables). In this context, various conditions for curvatures on Kähler manifolds under which they become Stein manifolds have been found [20].

Let H be complete and simply connected, and $K \leq 0$. Busemann functions on H are differentiable of class C^2 , and the horosphere $F_\gamma^{-1}(\{t\})$ is a C^2 -surface [22]. On a parabolic visibility manifold (\rightarrow Section F(3)) the negative of every Busemann function is C^1 -convex without minimum [7].

(2) Ends and Splitting Theorems. Ends of a noncompact M are defined as follows: If A_1 and A_2 are compact subsets of M with $A_1 \subset A_2$, then any component of $M - A_2$ is contained in a unique component of $M - A_1$. An **end** is the limit of an inverse system $\{\text{components of } M - A; A\}$ directed by the inclusion relation as indicated above and indexed by $\{A \mid A \text{ compact}\}$.

It has been shown that if $\text{Ric} > 0$ [23] (or ≥ 0 [19]), then M has exactly one end (or at most two ends). A visibility manifold has at most two ends if it is not Fuchsian [7]. If M admits a locally nonconstant convex function, then M has at most two ends [24].

If M has more than one end, then there exists a straight line $\gamma: \mathbf{R} \rightarrow M$ which is by definition a nontrivial geodesic any of whose subarcs is minimizing. A classical result of Cohn-Vossen (*Mat. Sb.*, 1 (1936)) states that if $\dim M = 2$, $K \geq 0$ and if M contains a straight

line, then the total curvature is 0 and hence M is isometric to either a plane or a cylinder $S^1 \times \mathbf{R}$. Generalizations of this result state that if M admits k independent straight lines and if $K \geq 0$ (Toponogov, *Amer. Math. Soc. Transl.*, 37 (1964)) (or if $\text{Ric} \geq 0$ [19]), then M is isometric to the Riemannian product $N \times \mathbf{R}^k$.

(3) Structure Theorems. The structure theorem of 2-dimensional M with $K > 0$ was proved by Cohn-Vossen (*Compositio Math.*, 2 (1935)). It has been shown that M is diffeomorphic to \mathbf{R}^m if $K > 0$ [23] and that if $K \geq 0$, then there exists a compact totally geodesic submanifold S without boundary such that M is homeomorphic to the total space of the †normal bundle over S [21]. Here, homeomorphism can be replaced by diffeomorphism, and if $K > 0$ outside a compact set of M , then M is diffeomorphic to \mathbf{R}^m [26].

For a ray γ and a point x on H , if σ is a ray with $\sigma(0) = x$ and if $d(\gamma(t), \sigma(t))$, $t > 0$, is bounded above, then σ is called an **asymptotic ray** (a ray asymptotic) to γ . The asymptotic relation is an equivalence relation on the set of all geodesics on H . A **point at infinity** of H is an equivalence class on geodesics on H , and the set of all points at infinity of H is denoted by $H(\infty)$. With a suitable topology, the set of all points at infinity of H constitutes a bounding sphere such that $\bar{H} = H \cup H(\infty)$ is a closed m -cell. For a properly discontinuous group D of isometries acting on H , a closed D -invariant limit set $L(D)$ is obtained in $H(\infty)$ as the set of all accumulation points of $\delta(p)$, $\delta \in D$. $M = H/D$ is a **visibility manifold** if and only if H satisfies the following axiom: any two distinct points on $H(\infty)$ can be joined by at least one geodesic. If $K \leq c < 0$ is satisfied on M , then M is a visibility manifold [7]. By investigating limit sets, visibility manifolds are classified into three types [7]. (1) M is parabolic; M is diffeomorphic to $N \times \mathbf{R}$ and is characterized by the fact that it has a convex function without minimum. (2) M is axial; M is a vector bundle over S^1 , and hence it is diffeomorphic to either $S^1 \times \mathbf{R}^{m-1}$ or the product of a Möbius strip with \mathbf{R}^{m-2} . (3) M is Fuchsian; M has more than two ends, and a strong algebraic restriction is imposed on the fundamental group of M .

(4) Convex Functions. Some elementary properties of C^∞ convex functions on M have been investigated and it was proved in [4] that if M admits a C^∞ convex function without minimum and if $K \leq 0$, then M is diffeomorphic to $N \times \mathbf{R}$, where N is a level hypersurface. When $K > 0$, F can be replaced by a strongly convex exhaustion function. Namely, for every compact set $A \subset M$ there is a $\delta > 0$

such that the second difference quotient along every geodesic at any point on A is bounded below by δ [20]. A standard †convolution smoothing procedure yields smooth convex approximations in neighborhoods of compact sets for every strongly convex function such that their second derivatives along every geodesic is positive. Global approximations can be constructed [27]. Let $\varphi: M \rightarrow \mathbf{R}$ be a convex function which is not constant on any open set. It has recently been proved that (1) if $b > \inf_M \varphi \geq -\infty$, then there is a homeomorphism $H: \varphi^{-1}(\{b\}) \times (\inf_M \varphi, \infty) \rightarrow M - \{\text{minimum set of } \varphi, \text{ if any}\}$ such that $\varphi \circ H(y, \alpha) = \alpha$ for all $y \in \varphi^{-1}(\{b\})$ and for all $\alpha \in (\inf_M \varphi, \infty)$; (2) if φ takes a minimum, then the continuous extension $\bar{H}: [\inf_M \varphi, \infty) \rightarrow M$ of H is proper and surjective [24].

G. Manifolds All of Whose Geodesics Are Closed

As is well known, every symmetric space of compact type of rank 1 (it is abbreviated CROSS as in [28]) has the property that all geodesics are simply closed and of the same length. Let M be a compact Riemannian manifold with the property that all geodesics on M are simply closed and that they have the same length (SC property). The problem discussed here is whether such an M is isometric (or at least the topology of such an M is equivalent) to some CROSS, or if it is possible to classify all such manifolds.

An example of a nontrivial SC-manifold was first constructed by Zoll (*Math. Ann.*, 57 (1903)) on S^2 as a surface of revolution in \mathbf{R}^3 , which is not isometric to a standard sphere. Blaschke conjectured that every SC-structure on $\mathbf{P}\mathbf{R}^2$ is the standard real projective space. This has been solved affirmatively by L. W. Green (*Ann. Math.*, 78 (1963)). In higher dimensions it has been proved that every infinitesimal deformation of the standard SC-structure on $\mathbf{P}\mathbf{R}^m$ is trivial [29]. A general result for a SC-manifold M states that the volume of M (with $\dim M = m$) with period 2π is the integral multiple of the volume of the standard unit m -sphere and this integer is a topological invariant [30]. For a point $p \in M$, if every geodesic segment with length l emanating from p is a simple geodesic loop at p , then M is called a **SC^p-manifold**. It has been proved that the †integral cohomology ring of every SC^p-manifold is isomorphic to one of the CROSS [31], which is a generalization of Bott's theorem (*Ann. Math.*, 60 (1954)).

An essential difference between the metrics of a Zoll surface and a standard sphere is seen

on the cut locus and the conjugate locus of a point. The (tangent) cut locus of any fixed point of a CROSS coincides with the (tangent) first conjugate locus; however, this does not hold on a Zoll surface. This observation gives rise to the definition of a Blaschke manifold: For distinct points $p, q \in M$, let $\Lambda(p, q) \subset M_q$ be the set of all unit vectors tangent to the minimizing geodesics from q to p . M is said to be a **Blaschke manifold at a point p** if for every $q \in C(p)$, $\Lambda(p, q)$ is a great sphere of the unit hypersphere in M_q centered at the origin. M is called a **Blaschke manifold** if it is so at every point of it. The following statements are equivalent [28]: (1) M is a Blaschke manifold at p , (2) the tangent cut locus C_p at p is spherical in M_p , (3) along every geodesic starting from p , the first conjugate point p appears at a constant length and the multiplicity of it is independent of the initial direction. Then such an M can be exhibited as $D \cup_a E$, where D is a closed ball with center p , E a C^∞ -closed k -disk bundle over an $(m-k)$ -dimensional C^∞ -compact manifold with boundary ∂E diffeomorphic to S^{m-1} , and $a: \partial D \rightarrow \partial E$ an attaching diffeomorphism. Conversely, if M is exhibited as $D \cup_a E$ then there exists a metric g on M such that (M, g) is a pointed Blaschke manifold at p , where p is the center of D . A generalized Blaschke's conjecture states that every Blaschke manifold is isometric to a CROSS. A partial solution to this conjecture has been obtained by M. Berger and states that if (S^m, g) is a Blaschke manifold, then it is isometric to a standard sphere [28].

H. G -Spaces

G -spaces were created by Busemann to show that many global theorems of differential geometry are independent of the Riemannian character of the metric and also of smoothness. This approach naturally led to novel questions and results. (Many facts hold even when the distance is not symmetric; \rightarrow [34].) The symbol G suggests that the principal property of G -spaces is the existence of geodesics with the same properties as in complete Riemannian manifolds without boundary except for differentiability. A **G -space X** is defined as follows: (1) X is metric with (symmetric) distance d ; (2) every bounded infinite set has at least one point of accumulation; (3) given two distinct points $p, r \in X$, there is a point $q \in X$ such that $p \neq q \neq r$ and $d(p, q) + d(q, r) = d(p, r)$; (4) to each point $x \in X$ a positive number ρ_x is assigned such that for any $p, q \in X$ with $d(p, x) < \rho_x$, $d(q, x) < \rho_x$ there is a point r with $d(p, q) + d(q, r) = d(p, r)$; (5) if

$d(p, q) + d(q, r) = d(p, r)$ and if $d(p, q) + d(q, r') = d(p, r')$, then $d(q, r) = d(q, r')$ implies $r = r'$.

From (2) and (3) it follows that any two points on X can be joined by a curve called a **segment** whose length realizes the distance between them. A **geodesic arc** (or for simplicity **geodesic**) $\gamma: [a, b] \rightarrow X$ is a curve such that any subarc contained in a ρ_x -ball around some point $x \in X$ is a segment. From (4) and (5) it follows that every geodesic arc has a unique infinite extension to both sides. If a geodesic arc is extended infinitely to both sides, then it is called a **geodesic line**.

The absence of smoothness causes an essential difference in the fact that the distance function to a fixed point on X is not necessarily convex in a small ball around it, in contrast to the Whitehead theorem (\rightarrow Section A) for Riemannian manifolds (or 'Finsler spaces). Here a function on X is called **convex** if and only if its restriction to every arc length-parametrized geodesic line is convex.

Every two points on X have neighborhoods which are homeomorphic to each other. Thus the topological dimension of X (\rightarrow 117 Dimension Theory) is well defined. A 1-dimensional X is either a circle S^1 or the whole real line. If $\dim X = 2$ [33] or if $\dim X = 3$ [36], then X is a topological manifold.

It should be noted that G -space theory not only proves known Riemannian theorems under weaker conditions, but also leads to many facts which were either not considered previously or well understood in (or thought to be very different from) the Riemannian case. Among many fruitful topics discussed by Busemann [33], only two are included below.

(1) **G -Surfaces.** For each point p on a 2-dimensional G -space \mathcal{S} , which is called a **G -surface**, let $S_p = \{x \in \mathcal{S} \mid d(x, p) = \rho\}$, where $0 < \rho < \rho_p/4$ is a fixed number. S_p is homeomorphic to a circle and bounds a 2-disk. An **angle A** at p is the set of all segments emanating from p and passing through the points on a connected subarc of S_p . An angular measure $|\cdot|$ for the angles at p is a nonnegative function which satisfies: (1) $|A| = \pi$ if and only if p is the midpoint of the segment joining two points on S_p which bounds A ; (2) if the intersection of two angles A_1 and A_2 is a unique segment, then $|A_1 \cup A_2| = |A_1| + |A_2|$. If a triple of points (p_0, p_1, p_2) on \mathcal{S} are contained in a sufficiently small ball, then the three segments joining them define a **triangle**, and each vertex p_i has the angle $|p_{i-1}p_i p_{i+1}|$ determined by the two segments. An **excess** $\varepsilon(p_0 p_1 p_2)$ of a triangle (p_0, p_1, p_2) is defined to be $\varepsilon(p_0 p_1 p_2) = |p_0 p_1 p_2| + |p_1 p_2 p_0| + |p_2 p_0 p_1| - \pi$. A degenerate triangle has excess 0. If a triangle (a, b, c) , which consists of geodesic arcs, is simplicially decomposed

into a finite number of triangles (a_i, b_i, c_i) , $i = 1, \dots, k$, then $\varepsilon(abc) = \sum_{i=1}^k \varepsilon(a_i b_i c_i)$, and this is independent of the choice of finite decomposition. If \mathcal{S} is compact and orientable, the **total excess** $\varepsilon(\mathcal{S})$ of \mathcal{S} is defined by $\varepsilon(\mathcal{S}) = \sum_{i=1}^k \varepsilon(a_i b_i c_i)$, where \mathcal{S} is simplicially decomposed into nondegenerate finite triangles (a_i, b_i, c_i) , $i = 1, \dots, k$. Then $\varepsilon(\mathcal{S}) = 2\pi\chi(\mathcal{S})$, where $\chi(\mathcal{S})$ is the Euler characteristic of \mathcal{S} . The results obtained by Cohn-Vossen [35] on the total curvature of complete open surfaces have also been generalized to G -surfaces with angular measure uniform at π . Although no angular measure leading to a true analog of the Gauss-Bonnet theorem exists even in G -surfaces [34], many of its applications, in particular the results by Cohn-Vossen, can be proved by using angular measures which are uniform at π [33].

If every two points on X can be joined by a unique geodesic line, then all geodesic lines on X are simultaneously either straight lines or circles of the same length l . If in the latter case $\dim X > 1$, then X has a two-fold universal covering space \tilde{X} whose geodesic lines are all circles of the same length $2l$, and they have the property that all geodesics through a point on \tilde{X} meet at the same point at length l . For the Riemannian case it is a Blaschke manifold (\rightarrow Section G). If every two points on \mathcal{S} can be joined by a unique geodesic line, then \mathcal{S} is either a plane or a projective plane.

(2) G -Spaces with Nonpositive Curvature. Let (p_0, p_1, p_2) be a triple of points on X which are contained in a small ball. X , by definition, is of **nonpositive curvature** if and only if for any such triple of points

$$d(p_1, p_{i+1}) \geq 2d(p'_i, p_{i+1}) \quad \text{for } i=0, 1, 2, \quad (*)$$

where p'_i is the midpoint of the unique segment joining p_{i-1} to p_{i+1} . If the above inequality is strict for any nondegenerate small triangle, then X is said to be of **negative curvature**. X of curvature 0 is defined in the same way. It should be noted that a G -space of curvature 0 is a locally Minkowskian space [33]. The sectional curvature of a Riemannian manifold M is nonpositive if and only if (*) holds for all small triangles on it. Because of $d(p_0, p'_0) \leq d(p_0, p'_1) + d(p'_1, p'_0) \leq \{d(p_0, p_1) + d(p_0, p_2)\}/2$, the distance function to a point $x \in X$ of nonpositive curvature is convex on a small ball around x . X is called **straight** if and only if all nontrivial geodesic lines on X are straight lines. If X has nonpositive curvature, then the universal covering space \tilde{X} is straight. Moreover, for any two geodesic lines $\gamma, \sigma: \mathbf{R} \rightarrow X$, the function $t \rightarrow d(\sigma(t), \gamma(t))$ is convex. In particular, the distance function to every fixed point on \tilde{X} is convex, and (*) holds for any

triangle on \tilde{X} . Thus \tilde{X} is contractible. Every nontrivial element of the homotopy class of loops on X with base point $x \in X$ contains a geodesic loop at x which has the minimum length. This fact and the strict convexity of $t \rightarrow d(\sigma(t), \gamma(t))$ for any two arc length parametrized geodesic lines γ, σ on the universal covering \tilde{X} of X with negative curvature imply that if X is compact, then every Abelian subgroup of the fundamental group $\pi_1(X)$ of X is infinite cyclic. This is a generalization of the Preissmann theorem (\rightarrow Section B).

References

- [1] A. D. Alexandrov, *Die innere Geometrie der konvexen Flächen*, Akademie-Verlag, 1955.
- [2] J. Cheeger and D. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland, 1975.
- [3] D. Gromoll, W. Klingenberg, and W. Meyer, *Riemannsche Geometrie in Grossen*, Lecture notes in math. 55, Springer, 1968.
- [4] R. Bishop and B. O'Neill, *Manifolds of negative curvature*, Trans. Amer. Math. Soc., 145 (1969), 1–49.
- [5] D. Gromoll and J. Wolf, *Some relations between the metric structure and the algebraic structure of the fundamental group in manifolds of nonpositive curvature*, Bull. Amer. Math. Soc., 77 (1971), 545–552.
- [6] B. Lawson and S. T. Yau, *On compact manifolds of nonnegative curvature*, J. Differential Geometry, 7 (1972), 211–228.
- [7] P. Eberlein and B. O'Neill, *Visibility manifolds*, Pacific J. Math., 46 (1973), 45–109.
- [8] N. Wallach, *Homogeneous Riemannian manifolds of strictly positive curvature*, Ann. Math., (2) 75 (1972), 277–295.
- [9] T. Sakai, *Cut loci of Berger's sphere*, Hokkaido Math. J., 10 (1981), 143–155.
- [10] J. Cheeger, *Finiteness theorems for Riemannian manifolds*, Amer. J. Math., 92 (1970), 61–74.
- [11] A. Weinstein, *On the homotopy type of positively pinched manifolds*, Arch. Math., 18 (1967), 523–524.
- [12] M. Gromov, *Manifolds of negative curvature*, J. Differential Geometry, 13 (1978), 223–230.
- [13] M. Gromov, *Almost flat manifolds*, J. Differential Geometry, 13 (1978), 231–241.
- [14] D. Gromoll, *Differenzierbare Struktur und Metriken positiver Krümmung auf Sphären*, Math. Ann., 164 (1966), 353–371.
- [15] Y. Shikata, *On the differentiable pinching problem*, Osaka J. Math., 4 (1967), 279–287.
- [16] M. Sugimoto and K. Shiohama, *On the*

differentiable pinching problem, *Math. Ann.*, 195 (1971), 1–16.

[17] E. Ruh, Curvature and differentiable structures on spheres, *Comment Math. Helv.*, 46 (1971), 127–136.

[18] K. Grove and K. Shiohama, A generalized sphere theorem, *Ann. Math.*, (2) 106 (1977), 201–211.

[19] J. Cheeger and D. Gromoll, The splitting theorem for manifolds of nonnegative Ricci curvature, *J. Differential Geometry*, 6 (1971), 119–129.

[20] R. Greene and H. Wu, Function theory on manifolds which possess a pole, *Lecture notes in math.* 699, Springer, 1979.

[21] J. Cheeger and D. Gromoll, On the structure of complete manifolds of nonnegative curvature, *Ann. Math.*, (2) 96 (1972), 413–443.

[22] H. Im Hof and E. Heitze, Geometry of horospheres, *J. Differential Geometry*, 12 (1977), 481–492.

[23] D. Gromoll and W. Meyer, On complete open manifolds of positive curvature, *Ann. Math.*, (2) 90 (1969), 75–90.

[24] R. Greene and K. Shiohama, Convex functions on complete noncompact manifolds; Topological structures, *Inventiones Math.*, 63 (1981), 129–157.

[25] W. Poor, Some results on nonnegatively curved manifolds, *J. Differential Geometry*, 9 (1974), 583–600.

[26] H. Wu, An elementary proof in the study of nonnegative curvature, *Acta Math.*, 142 (1979), 57–78.

[27] R. Greene and H. Wu, C^∞ convex functions and manifolds of positive curvature, *Acta Math.*, 137 (1976), 209–245.

[28] A. Besse, *Manifolds all of whose geodesics are closed*, Springer, 1978.

[29] R. Michel, Problème d'analyse géométrique lies à la conjecture de Blaschke, *Bull. Soc. Math. France*, 101 (1973), 17–69.

[30] A. Weinstein, On the volume of manifolds whose geodesics are all closed, *J. Differential Geometry*, 9 (1974), 513–517.

[31] H. Nakagawa, A note on theorems of Bott and Samelson, *J. Math. Kyoto Univ.*, 7 (1967), 205–220.

[32] C. T. Yang, Odd-dimensional Wiederschen manifolds are spheres, *J. Differential Geometry*, 15 (1980), 91–96.

[33] H. Busemann, *The geometry of geodesics*, Academic Press, 1955.

[34] H. Busemann, *Recent synthetic differential geometry*, Springer, 1970.

[35] S. Cohn-Vossen, *Kürzeste Wege und Totalkrümmung auf Flächen*, *Compositio Math.*, 2 (1935), 63–133.

[36] B. Krakus, Any 3-dimensional G -space is a manifold, *Bull. Acad. Pol. Sci.*, 16 (1968), 737–740.

179 (VI.5) Geometric Construction

A. General Remarks

A **geometric construction problem** is a problem of drawing a figure satisfying given conditions by using certain prescribed tools only a finite number of times. If the problem is solvable, then it is called a **possible construction problem**; if it is unsolvable, even though there exist figures satisfying the given conditions, then it is an **impossible construction problem**. If there does not exist a figure satisfying the given conditions, then we say that the problem is **inconsistent**.

Among problems of geometric construction, the oldest and the best known are those of constructing plane figures by means of **ruler** and **compass**. In this article, we call these problems simply problems of elementary geometric construction. The following are some of the more famous problems of this kind; the first four are possible construction problems.

(1) Suppose that we are given three straight lines l, m, n and three points P, Q, R in a plane. Draw a triangle ABC in such a way that vertices A, B, C lie on l, m, n and sides BC, CA, AB pass through P, Q, R (**Steiner's problem**).

(2) Suppose that we are given a circle O and three points P, Q, R not lying on O . Draw a triangle ABC inscribed in O in such a way that the sides BC, CA, AB pass through P, Q, R (**Cramer-Castillon problem**).

(3) Draw a circle tangent to all of three given circles (**Apollonius' problem**).

(4) Suppose that we are given a triangle. Draw three circles inside this triangle in such a way that each is tangent to two sides of the triangle and any two of the circles are tangent to each other (**Malfatti's problem**).

(5) Let n be a natural number. For the division of the circumference of a circle into n equal parts (consequently, the construction of a n -gon) to be a possible construction problem, it is necessary and sufficient that the representation of n as a product of prime numbers take the form $n = 2^\lambda p_1 \dots p_k$, where $\lambda \geq 0$, p_1, \dots, p_k are all different prime numbers of the form $2^h + 1$ (h -Fermat number) (C. F. Gauss, 1801).

(6) The following are three famous impossible construction problems of Greek origin: (i) divide a given angle into three equal parts (**trisection of an angle**); (ii) construct a cube whose volume is double that of a given cube (**duplication of a cube** or the **Delos problem**); and (iii) construct a square whose area is that of a given circle (**quadrature of a circle**). P. L. Wantzel (1837) proved that problems (i) and

(ii) are impossible except for the special cases of $\pi/2$, $\pi/4$, etc. for (i); and C. L. F. Lindemann (1882) proved the impossibility of (iii) while proving that the number π is \dagger transcendental.

B. Conditions for Constructibility

A problem of elementary geometric construction amounts to a problem of determining a certain number of points by drawing straight lines that pass through given pairs of points, and circles having given points as centers and passing through given points. Let (a_1, b_1) , (a_2, b_2) , ..., (a_n, b_n) be rectangular coordinates of given points, and let K be the smallest \dagger number field containing the numbers a_1 , ..., b_n . Straight lines that join given pairs of points and circles that have given points as centers and that pass through given points are represented by equations of the first or second degree with coefficients belonging to K . Consequently, the coordinates of points of intersection of these straight lines and circles belong to a quadratic extension $K' = K(\sqrt{d})$ of K . Let A be the set of coordinates of the points that are to be determined. Then the problem is solvable if and only if any number α in A is contained in a field $L = K(\sqrt{d_1}, \sqrt{d_2}, \dots, \sqrt{d_r})$, where $d_{i+1} \in K(\sqrt{d_1}, \dots, \sqrt{d_i})$ ($i = 0, 1, \dots, r-1$). Thus L is a \dagger normal extension field of K whose degree over K is a power of 2. Using this theorem we can prove the impossibility of trisection of an angle and duplication of a cube.

Since the 18th century, besides the problem of construction by ruler and compass, problems of construction by ruler alone or by compass alone have also been studied. We state here some of the more notable results: (1) If by drawing a straight line we mean the process of finding two different points on that line, then we can solve all the problems of elementary geometric construction by means of compass alone (G. Mohr, L. Mascheroni). (2) If by drawing a circle we mean the process of finding its center and a point on its circumference, and if a circle and its center are given, then we can solve any problem of elementary geometric construction by means of ruler alone. (3) It is not possible to find the center of a given circle by ruler alone (D. Hilbert). (4) It is impossible to bisect a given segment by ruler alone. (5) When two intersecting circles or concentric circles are given, we can find the centers of these circles by ruler alone. When nonintersecting and nonconcentric circles are given, it is not possible to find their centers by ruler alone (D. Cauer).

Cases have been considered in which the radius of a circle we can draw by compass or

the length of a segment we can draw by ruler is required to satisfy certain conditions. Also, various considerations have been made concerning cases in which we can use tools other than ruler and compass. For example, it is known that although not all possible elementary geometric construction problems are solvable by ruler alone, all these problems are possible if we have either a pair of parallel rulers, a square, or a triangle having a fixed acute angle. If we use a square and a compass, then the trisection of an angle and the duplication of a cube are possible (L. Bieberbach). Also, when a conic section other than a circle is given, the trisection of an angle and the duplication of a cube become possible by ruler and compass (H. J. S. Smith and H. Kortum). By ruler and \dagger transfer of constant lengths, we can solve Malfatti's problem but not Apollonius' problem (Feldblum).

Even when a problem is possible, the method of construction may be rather complicated and impractical. In these cases, various methods of highly accurate approximate construction have been investigated.

References

- [1] H. L. Lebesgue, *Leçons sur les constructions géométriques*, Gauthier-Villars, 1950.
- [2] L. Bieberbach, *Theorie der geometrischen Konstruktionen*, Birkhäuser, 1952.

180 (XX.15) Geometric Optics

A. General Remarks

Geometric optics is a mathematical theory of light rays. It is not concerned with the properties of light rays as waves (e.g., their wavelength and frequency), but studies their properties as pencils of rays that follow three laws: the law of rectilinear propagation, the law of reflection (i.e., angles of incidence and reflection on a smooth plane are equal (Euclid)), and the law of refraction (i.e., if θ and θ' are angles of incidence and refraction of a light ray refracted from a uniform medium to a second uniform medium and if n , n' are the refractive indices of the first and the second medium, respectively, then $n \sin \theta = n' \sin \theta'$ (R. W. Snell, Descartes)). These three laws follow from **Fermat's principle**, which states that the path of a light ray traveling from a point A' to A in a medium with refractive index $n(P)$ at P is such that the integral $\int_A^{A'} n(P) ds$ attains its

extremal value, where ds is the line element along the path. This line integral is called the **optimal distance** from A' to A . Therefore Fermat's principle can be taken as a foundation of geometric optics and is, in a way, similar to the \dagger variational principle in particle dynamics (**Maupertuis's principle**),

$$\delta \int \sqrt{2h - 2U(P)} ds = 0,$$

which is satisfied by the path of a particle of unit mass having constant energy h passing through a field of \dagger potential $U(P)$. The quantity $\sqrt{2h - 2U}$ corresponds to the refractive index n .

In an optical system, express the position of a point on the path of a light ray by orthogonal coordinates (x, y, z) , and define the \dagger Lagrangian $L = n\sqrt{1 + \dot{x}^2 + \dot{y}^2}$ ($\dot{x} = dx/dz$, $\dot{y} = dy/dz$), **optical direction cosines** $p = \partial L / \partial \dot{x}$, $q = \partial L / \partial \dot{y}$, and the \dagger Hamiltonian $H = \dot{x}p + \dot{y}q - L = -\sqrt{n^2 - p^2 - q^2}$. Then the \dagger canonical equations of the path are obtained as in particle dynamics; x, y and p, q are called \dagger canonical variables. Due to the variational principle, the integral of the linear differential form $p dx + q dy - H dz = \omega_d$ along a light path is a function $S(A', A)$ of the endpoints A', A of the path, and the optical direction cosines and the Hamiltonians of the system at A and A' are given by

$$\frac{\partial S}{\partial x'} = -p', \quad \frac{\partial S}{\partial y'} = -q', \quad \frac{\partial S}{\partial z'} = H',$$

$$\frac{\partial S}{\partial x} = p, \quad \frac{\partial S}{\partial y} = q, \quad \frac{\partial S}{\partial z} = -H.$$

Hence we obtain \dagger Hamilton-Jacobi differential equations

$$\left(\frac{\partial S}{\partial x'}\right)^2 + \left(\frac{\partial S}{\partial y'}\right)^2 + \left(\frac{\partial S}{\partial z'}\right)^2 = n'^2,$$

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = n^2.$$

As a corollary to these relations we obtain **Malus's theorem**, which states that a pencil of light rays perpendicular to a common surface (locally) at a given moment is also perpendicular to a common surface (locally) after an arbitrary number of reflections and refractions.

Suppose that light rays travel from an object space into an image space through an optical apparatus. If all the rays starting from any one point of the object space converge to a point of the image space and if the mapping given by this correspondence is \dagger bijective, then we say that this imaging is **perfect**. Examples of perfect imaging systems are realized by optical apparatus such as **Maxwell's fisheye** (having refractive index $n(r) = a/(b + r^2)$, where

r denotes the distance from the center of the system) and **Luneburg's lens** ($n(r) = \sqrt{a - r^2}$).

Perfect imaging conserves optical distance, yields the relation $n(A')ds' = n(A)ds$, and gives a \dagger conformal mapping, with the magnification inversely proportional to the refractive index.

B. Gauss Mappings

Consider an optimal system with a symmetrical axis of rotation, its **optical axis**. A ray of light that is near the optical axis and has a small inclination to the axis is called a **paraxial ray**. A mapping realizable by paraxial rays where the canonical variables x, y, p, q can be considered to be infinitesimal variables whose squares are negligible, is called a **Gauss mapping** (**Gauss map**). When the positions of an object point and its image under a Gauss mapping are represented by homogeneous coordinates, the mapping is represented as a linear transformation, i.e., a \dagger collineation, which maps a point to a point and a line to a line. A point in one space corresponding to the point at infinity in the other space is called a **focus**. If we take a focus as the origin of a coordinate system in each space and use the \dagger homogeneous coordinates x_i such that $x = x_1/x_4$, $y = x_2/x_4$, $z = x_3/x_4$, then a Gauss mapping can be represented as $x_1 = x'_1$, $x_2 = x'_2$, $x_3 = f x'_4$, $f' x'_4 = x'_3$. The ratio of x to x' , i.e., the lateral magnification, is $x/x' = z/f = f'/z'$, where x' is the length of an object orthogonal to the axis and x is the length of its image. The distance f between a focus and a point where the lateral magnification is 1 (such a point is called a **principal point**) is called **focal length** in each space. The telescopic mapping, i.e., $x_1 = x'_1$, $x_2 = x'_2$, $x_3 = a x'_3$, $x_4 = b x'_4$, is also a Gauss mapping, in which the lateral magnification is constant.

C. Aberration

When a mapping is realized not only by paraxial rays but also by rays having larger inclinations, a departure from the Gauss mapping arises. This departure is generally called **aberration**. Suppose that a light ray that passes through the point (x', y', z') of a plane perpendicular to the optical axis at a fixed z' and has optical direction cosines p', q' is transformed by the optical apparatus into a light ray that passes through the point (x, y, z) of a plane perpendicular to the optical axis at a fixed z and has optical direction cosines p, q there. Then by the variational principle, $p dx + q dy - p' dx' - q' dy' = dW$ (dW is an \dagger exact differential). Therefore the transformation $(x', y', p', q') \rightarrow (x, y, p, q)$ is a \dagger canonical transformation. The

mapping can be described by

$$p = \frac{\partial W}{\partial x}, \quad q = \frac{\partial W}{\partial y},$$
$$p' = -\frac{\partial W}{\partial x'}, \quad q' = -\frac{\partial W}{\partial y'}$$

in terms of W and can also be represented in terms of $V = W + p'x' + q'y'$ or $U = W + p'x' + q'y' - px - qy$. For a given optical system, one of these functions W, U, V (called a **characteristic function** or **eikonal**) can be used to estimate the aberration. By developing such a characteristic function in power series of canonical variables and observing its terms of less than the fifth power, we can single out five kinds of aberration: spherical aberration, curvature of image field, distortion, coma, and astigmatism in a rotationally symmetric optical system. To eliminate these aberrations an optical system must satisfy Abbe's sine condition (the elimination of spherical aberration and coma), Petzval's condition (the elimination of astigmatism and curvature of image field), and the tangent condition (the elimination of distortion).

The path of a charged particle in an electromagnetic field can be treated in the same way as the path of a light ray. Let ϵ represent the specific charge of the particle, h the energy, A_0 the electrostatic potential, and A_x, A_y, A_z vector potentials. Then the index of refraction is $\sqrt{2(h - \epsilon A_0) + \epsilon(A_x dx/ds + A_y dy/ds + A_z dz/ds)}$. In this case, the index of refraction shows the anisotropy caused by the existence of the magnetic field. The paths of paraxial rays are determined by a set of linear differential equations of the second order, and the Gauss mapping is realized as in geometric optics.

References

[1] C. Carathéodory, *Geometrische Optik*, Springer, 1937.
[2] M. Herzberger, *Modern geometrical optics*, Interscience, 1958.
[3] R. K. Luneburg, *Mathematical theory of optics*, Univ. of California Press, 1964.
[4] O. N. Stavroudis, *The optics of rays, wavefronts, and caustics*, Academic Press, 1972.

181 (VI.1)
Geometry

The Greek word for geometry, which means *measurement of the earth*, was used by the

historian Herodotus, who wrote that in ancient Egypt people used geometry to restore their land after the inundation of the Nile. Thus the theoretical use of figures for practical purposes goes back to pre-Greek antiquity. Tradition holds that Thales of Miletus knew some properties of congruent triangles and used them for indirect measurement, and that the Pythagoreans had the idea of systematizing this knowledge by means of proofs (\rightarrow 24 Ancient Mathematics; 187 Greek Mathematics). \dagger Euclid's *Elements* is an outgrowth of this idea [1]. In this work, we can see the entire mathematical knowledge of the time presented as a logical system. It includes a chapter (Book V) on the theory of quantity (i.e., the theory of positive real numbers in present-day terminology) and chapters on the theory of integers (Books VII–IX), but for the most part, it treats figures in a plane or in space and presents number-theoretic facts in geometric language.

Geometry in today's usage means the branch of mathematics dealing with spatial figures. In ancient Greece, however, all of mathematics was regarded as geometry. In later times, the French word *géomètre* or the German word *Geometer* was sometimes used as a synonym for *mathematician*. In a fragment of his *Pensées*, B. Pascal speaks of the *esprit de géométrie* as opposed to the *esprit de finesse*. The former means simply the mathematical way of thinking.

Algebra was introduced into Europe from the Middle East toward the end of the Middle Ages and was further developed during the Renaissance. In the 17th and the 18th centuries, with the development of analysis, geometry achieved parity with algebra and analysis.

As R. Descartes pointed out, however, figures and numbers are closely related [2]. Geometric figures can be treated algebraically or analytically by means of \dagger coordinates (the method of **analytic geometry**, so named by S. F. Lacroix [3]); conversely, algebraic or analytic facts can be expressed geometrically. Analytic geometry was developed in the 18th century, especially by L. Euler [4], who for the first time established a complete algebraic theory of \dagger curves of the second order. Previously, these curves had been studied by Apollonius (262–200? B.C.) as \dagger conic sections. The idea of Descartes was fundamental to the development of analysis in the 18th century. Toward the end of that century, analysis was again applied to geometry. For example, G. Monge's contribution [5] can be regarded as a forerunner of \dagger differential geometry.

However, we cannot say that the analytic method is always the best manner of dealing with geometric problems. The method of treating figures directly without using coordinates

is called **synthetic** (or **pure**) **geometry**. In this vein, a new field called \dagger projective geometry was created by G. Desargues and B. Pascal in the 17th century. It was further developed in the 19th century by J.-V. Poncelet, L. N. Carnot, and others. In the same century, J. Steiner insisted on the importance of this field (\rightarrow 343 Projective Geometry).

On the other hand, the \dagger axiom of parallels in Euclid's *Elements* has been an object of criticism since ancient times. In the 19th century, by denying the a priori validity of Euclidean geometry, J. Bolyai and N. I. Lobachevskii formulated non-Euclidean geometry, whose logical consistency was shown by models constructed in both Euclidean and projective geometry (\rightarrow 285 Non-Euclidean Geometry).

In analytic geometry, physical spaces and planes, as we know them, are represented as 3-dimensional or 2-dimensional Euclidean spaces E^3 , E^2 . It is easy to generalize these spaces to n -dimensional Euclidean space E^n . A "point" of E^n is an n -tuple of real numbers (x_1, \dots, x_n) , and the distance between two points (x_1, \dots, x_n) , (y_1, \dots, y_n) is $((y_1 - x_1)^2 + \dots + (y_n - x_n)^2)^{1/2}$. The geometries of E^2 , E^3 are called **plane geometry** and **space** (or **solid**) **geometry**, respectively. The geometry of E^n is called **n -dimensional Euclidean geometry**. We obtain n -dimensional projective or non-Euclidean geometries similarly. F. Klein [7] proposed systematizing all these geometries in group-theoretic terms. He called a "space" a set S on which a group G operates and a "geometry" the study of properties of S invariant under the operations of G (\rightarrow 137 Erlangen Program).

B. Riemann [6] initiated another direction of geometric research when he investigated n -dimensional \dagger manifolds and, in particular, \dagger Riemannian manifolds and their geometries. Some aspects of Riemannian geometry fall outside of geometry in the sense of Klein. It was a starting point for the broad field of modern differential geometry, that is, the geometry of \dagger differentiable manifolds of various types (\rightarrow 109 Differential Geometry).

The reexamination of the system of axioms of Euclid's *Elements* led to D. Hilbert's \dagger foundations of geometry and to the axiomatic tendency of present-day mathematics. The study of algebraic curves, which started with the study of conic sections, developed into the theory of algebraic manifolds, the algebraic geometry that is now developing so rapidly (\rightarrow 12 Algebraic Geometry). Another branch of geometry is topology, which has developed since the end of the 19th century. Its influence on the whole of mathematics today is considerable (\rightarrow 114 Differential Topology; 426 Topology). Geometry has now permeated all

branches of mathematics, and it is sometimes difficult to distinguish it from algebra or analysis. The importance of geometric intuition, however, has not diminished from antiquity until today.

References

- [1] T. L. Heath, The thirteen books of Euclid's elements, Cambridge, 1908.
- [2] R. Descartes, Géométrie, Paris, 1637 (Oeuvres, IV, 1901).
- [3] S. F. Lacroix, Traité élémentaire de trigonométrie rectiligne et sphérique et d'application de l'algèbre à la géométrie, Bachelier, 1798–1799.
- [4] L. Euler, Introductio in analysin infinitorum, Lausanne, 1748 (Opera omnia VIII, IX, 1922); French translation, Introduction à l'analyse infinitésimale I, II, Paris, 1835; German translation, Einleitung in die Analysis des Unendlichen, Springer, 1885.
- [5] G. Monge, Application de l'analyse à la géométrie, fourth edition, Paris, 1809.
- [6] B. Riemann, Über die Hypothese, welche der Geometrie zu Grunde liegen, Habilitationsschrift, 1854 (Gesammelte mathematische Werke, Teubner, 1876, 254–269; Dover, 1953).
- [7] F. Klein, Vergleichende Betrachtungen über neuere geometrische Forschungen, Das Erlanger Programm, 1872 (Gesammelte mathematische Abhandlungen I, Springer, 1921, 460–497).

182 (V.10) Geometry of Numbers

A. History

H. Minkowski introduced the notions of lattice and convex set in the \dagger algebraic theory of numbers. He developed a simple yet powerful method of arithmetic investigation using these geometric notions to simplify the analytic theory of \dagger Diophantine approximation, which had been developed by P. G. L. Dirichlet and C. Hermite. His theory, the **geometry of numbers**, has continued its development and contributed to various fields of mathematics (\rightarrow 83 Continued Fractions).

B. Lattices

Let E^n be an n -dimensional Euclidean space identified with the linear space \mathbf{R}^n . For a point

P in E^n , we denote the corresponding vector in \mathbf{R}^n by $v(P) = \langle x_1, \dots, x_n \rangle$. A subset Λ of E^n is called an n -dimensional (**homogeneous**) **lattice** if there exists a basis $\{v_1, \dots, v_n\}$ of \mathbf{R}^n such that $\Lambda = \{P \in E \mid v(P) = \sum_{i=1}^n \lambda_i v_i, \lambda_i \in \mathbf{Z}\}$. The set of points $\{X_1, \dots, X_n\}$ such that $v(X_i) = v_i$ ($i = 1, \dots, n$) is called a **basis** of the lattice Λ . A typical example of a lattice is the point set corresponding to \mathbf{Z}^n in \mathbf{R}^n . The free module generated by v_i ($i = 1, \dots, n$) is denoted by Λ^* and is called the **lattice group** of Λ . We have $\Lambda^* = \{v \in \mathbf{R}^n \mid v = v(P), P \in \Lambda\}$. If $\{u_1, \dots, u_n\}$ is another basis of the free module Λ^* , then there exists an element (α_{ij}) of $GL(n, \mathbf{Z})$ (i.e., $\alpha_{ij} \in \mathbf{Z}$ and $|\det(\alpha_{ij})| = 1$) such that $u_j = \sum_{i=1}^n \alpha_{ij} v_i$. Hence the quantity $|\det(v_1, \dots, v_n)|$ is independent of the choice of the basis $\{v_1, \dots, v_n\}$. We denote this quantity by $d(\Lambda)$ and call it the **determinant of the lattice**. We denote the minimum distance between the points belonging to Λ by $\delta(\Lambda)$.

A subset L of the space E^n is called an **inhomogeneous lattice** if there exists a homogeneous lattice Λ in E^n and a point P_0 in E^n such that $L = \{P \in E^n \mid v(P) - v(P_0) \in \Lambda^*\}$. Thus an inhomogeneous lattice is obtained from a homogeneous lattice by translation. In this article we restrict ourselves to the case of homogeneous lattices and henceforth omit the adjective "homogeneous."

Suppose we are given a sequence of lattices $\Lambda_1, \Lambda_2, \dots$, in E^n with bases $\{X_1^{(1)}\}, \{X_1^{(2)}\}, \dots$. If the sequence of points $X_i^{(v)}$ converges to X_i ($i = 1, \dots, n$) and the set $\{X_1, \dots, X_n\}$ forms a basis of a lattice Λ , we call Λ the **limit of the sequence** $\{\Lambda_v\}$; we also say that the sequence $\{\Lambda_v\}$ **converges** to the lattice Λ . In this case we have $d(\Lambda_v) \rightarrow d(\Lambda)$, $\delta(\Lambda_v) \rightarrow \delta(\Lambda)$. The notion of convergence of lattices gives rise to a topology of the space M_0 of all the lattices in E . A sequence $\{\Lambda_v\}$ of lattices is said to be **bounded** if there exist positive numbers c and c' such that $d(\Lambda_v) \leq c$, $\delta(\Lambda_v) \geq c'$ for all v . A bounded sequence of lattices has a convergent subsequence.

Let S be a subset of the space E^n . A lattice Λ is called **S -admissible** if we have $\Lambda \cap S^i = \{O\}$, where S^i is the interior of S and O is the origin. We denote the set of S -admissible lattices by $A(S)$. Given a closed subset M of M_0 , we put $\Delta(S \setminus M) = \inf_{\Lambda \in A(S) \cap M} d(\Lambda)$ if $A(S) \cap M$ is non-empty, while if $A(S) \cap M$ is empty, we put $\Delta(S \setminus M) = \infty$. When $M = M_0$, we write $\Delta(S \setminus M) = \Delta(S)$ and call it the **critical determinant** of S . Generally, a lattice Λ in M is said to be **critical in M with respect to S** if $\Lambda \in A(S)$ and $d(\Lambda) = \Delta(S \setminus M)$. Suppose that we have $0 < \Delta(S \setminus M) < \infty$. Then for a lattice critical in M with respect to S to exist it is necessary and sufficient that there exist a bounded sequence $\{\Lambda_v\}$ such that $\Lambda_v \in M \cap A(S)$ and $d(\Lambda_v) \rightarrow \Delta(S \setminus M)$.

C. Successive Minima and Minkowski's Theorem

A subset S of the space E^n is called a **bounded star body** (symmetric with respect to the origin) if there exists a continuous function F defined on the space E satisfying the following four conditions: (i) $F(O) = 0$; (ii) if $X \neq O$, then $F(X) > 0$; (iii) for an arbitrary real number t and a point X , we have $F(tX) = |t|F(X)$; (iv) $S = \{X \mid F(X) \leq 1\}$. A bounded closed convex body that is symmetric with respect to the origin is a bounded star body. If we are given a star body S , the associated function F , and a lattice Λ , there exist a set of points $\{P_1, \dots, P_n\}$ in Λ and a set of positive numbers $\{p_1, \dots, p_n\}$ satisfying the following four conditions: (1) $v(P_1), \dots, v(P_n)$ are linearly independent; (2) $F(P_i) = p_i$ ($i = 1, \dots, n$); (3) $p_1 \leq \dots \leq p_n$; (4) if P is a point in Λ and $v(P)$ is not contained in the subspace spanned by $\{v(P_1), \dots, v(P_{n-1})\}$, then $F(P) \geq p_n$. The set $\{p_1, \dots, p_n\}$ is uniquely determined by S and Λ . The numbers p_i are called the **successive minima** of S in Λ ; the points P_i are the **successive minimum points** of S in Λ .

Minkowski's theorem: Let Λ be a lattice in a Euclidean space E^n and S a bounded subset of E^n . Then we have the following:

(I) If the volume $V(S)$ is larger than $d(\Lambda)$, then there exist points X_1 and X_2 in S such that $X_1 \neq X_2$ and $v(X_1) - v(X_2) \in \Lambda^*$. Suppose, moreover, that S is convex and symmetric with respect to the origin. Then, if $V(S) > 2^n d(\Lambda)$, there exists a point X in $S \cap \Lambda$ different from the origin. Hence we have $2^n \Delta(S) \geq V(S)$ ($n = \dim E^n$).

(II) Let S be a bounded closed convex body that is symmetric with respect to the origin, and let p_1, \dots, p_n be the successive minima of S in Λ . Then we have $p_1 \dots p_n \cdot V(S) \leq 2^n d(\Lambda)$.

D. Minkowski-Hlawka Theorem

Suppose that we are given a subset S of the n -dimensional Euclidean space E^n such that the characteristic function $\chi(X)$ of S is integrable in the sense of Riemann. Then we have the **Minkowski-Hlawka theorem**: (i) If $n \geq 2$ and S is open, then $\Delta(S) \leq V(S)$, and (ii) if, moreover, S is symmetric with respect to the origin, then $2\Delta(S) \leq V(S)$; (iii) if S is a symmetric star body with respect to the origin, then $2\zeta(n)\Delta(S) \leq V(S)$, where $\zeta(n)$ is the Riemann zeta function.

A proof for the theorem was given by E. Hlawka (1944); (iii) and (iv) were conjectured by Minkowski. C. L. Siegel obtained another proof (1945), and C. A. Rogers simplified the original proof by Hlawka (1947). There are

results concerning the estimation of $\Delta(S)/V(S)$ for various subsets S .

E. Siegel's Mean Value Theorem

In an attempt to obtain a proof for the latter half of the Minkowski-Hlawka theorem, Minkowski observed the necessity of establishing the arithmetic theory of the linear transformation groups. Siegel was inspired by this observation and obtained the following theorem, **Siegel's mean value theorem**, which implies the Minkowski-Hlawka theorem: Let F be a \dagger fundamental region of the group $SL(n, \mathbf{R})$ with respect to a discrete subgroup $SL(n, \mathbf{Z})$. Let ω be the \dagger invariant measure on $SL(n, \mathbf{R})$ such that $\int_F d\omega = 1$ (\rightarrow 225 Invariant Measures). Let f be a bounded Riemann integrable function with compact \dagger support defined on the space \mathbf{R}^n . Note that the lattice \mathbf{Z}^n is stabilized by the subgroup $SL(n, \mathbf{Z})$. We have

$$\int_F \sum_{\substack{x \in \mathbf{Z}^n \\ x \neq 0}} f(g \cdot x) d\omega(g) = \int_{\mathbf{R}^n} f(x) dx,$$

where the right-hand side of the equation is the usual Riemann integral of the function f . A. Weil considered this theorem in a more general setting (*Summa Brasil. Math.*, 1 (1946)).

F. Diophantine Approximation

Minkowski initiated the notion of Diophantine approximation in reference to the problem of estimating the absolute value $|f(x)|$ of a given function f , where x varies in \mathbf{Z} or in a given ring of \dagger algebraic integers. (A \dagger Diophantine equation is an equation $f(x)=0$, where x varies in \mathbf{Z} .) Today Diophantine approximation (in the wide sense) refers to the investigation of the scheme of values $f(x)$, where x varies in a suitable ring of algebraic integers. The geometry of lattices is a powerful tool in this investigation. A typical problem in this field of study is that of approximating irrational numbers by rational numbers; here \dagger continued fractions play an important role (\rightarrow Section G). For the problem of uniform distribution considered by H. Weyl, the analytic method, especially that of trigonometric series, is useful (\rightarrow Section H). **Dirichlet's drawer principle** (to put n objects in m drawers with $n > m$, it is necessary to put more than one object in at least one drawer) is one of the basic principles used in the theory of Diophantine approximation. Recently, the theory has been applied to the theory of \dagger transcendental numbers and the theory of \dagger Diophantine equations.

G. Approximation of Irrational Numbers by Rational Numbers

Given an irrational number θ , we have the problem of finding rational integers x (> 0) and y such that $|\theta - y/x| < \varepsilon/x$, where ε is a given positive number. Suppose that we are given a positive integer N . Using Dirichlet's drawer principle we can show the existence of x ($\leq N$) and y such that $|\theta - y/x| < 1/xN$. Let $M(\theta)$ be the supremum of positive numbers M such that the inequality $|\theta - y/x| < 1/Mx^2$ holds for infinitely many pairs of integers x, y . We have $1 \leq M(\theta) (\leq \infty)$. Two irrational numbers θ and θ' are said to be **equivalent** if there exists an element $(a_{ij}) \in GL(2, \mathbf{Z})$ such that $\theta' = (a_{11}\theta + a_{12})/(a_{21}\theta + a_{22})$. In this case we have $M(\theta) = M'(\theta)$.

If the irrational number θ satisfies the quadratic equation $a\theta^2 + b\theta + c = 0$ (a, b, c are rational integers), then we have $M(\theta) = k^{-1} \sqrt{b^2 - 4ac}$, where $k = \min\{ax^2 + bxy + cy^2 | x, y \in \mathbf{Z}, x \neq 0, y \neq 0\}$. In general, for an irrational number θ of degree two, we have $M(\theta) \geq \sqrt{5}$. The equality $M(\theta) = \sqrt{5}$ holds if θ is equivalent to $\theta_1 = (1 + \sqrt{5})/2$. If θ is not equivalent to θ_1 , then $M(\theta) \geq \sqrt{8}$; the equality holds if θ is equivalent to $\theta_2 = 1 + \sqrt{2}$. Similarly, we have $\theta_3, \theta_4, \dots$; and $M(\theta_n) \rightarrow 3$ ($n \rightarrow \infty$). If $M(\theta) < 3$, there exists a θ_n such that θ is equivalent to θ_n . The set of irrational numbers θ satisfying $M(\theta) = 3$ is uncountable (A. A. Markov [13]). We have no information about $M(\theta)$ for the general algebraic irrational number θ . Let $\mu(\theta)$ be the supremum of real numbers μ such that the inequality $|\theta - y/x| < 1/x^\mu$ holds for infinitely many pairs of integers x, y . Given a number $\kappa > 2$, it can be shown that the \dagger Lebesgue measure of the set of real numbers θ such that $\mu(\theta) \geq \kappa$ is zero. If θ is a real algebraic number of degree n , then $\mu(\theta) \leq n$ (J. Liouville). Concerning $\mu(\theta)$, results have been obtained by A. Thue, Siegel, A. O. Gel'fond, and F. J. Dyson. K. F. Roth (1954) proved that $\mu(\theta) = 2$ (**Roth's theorem** [12]), which settled the problem of $\mu(\theta)$. Roth's theorem means that if κ is larger than 2, then there exist only a finite number of pairs x, y satisfying $|\theta - y/x| < 1/x^\kappa$. This can be generalized to the case of the approximation of an element θ that is algebraic over an A -field k by an element of the field k (S. Lang [9]). (An A -field is either an algebraic number field of finite degree or an algebraic function field in one variable over a finite constant field.)

In 1970 W. M. Schmidt [20] obtained theorems on simultaneous approximation which generalize Roth's theorem. Thus, if $\alpha_1, \dots, \alpha_n$ are real algebraic numbers such that $1, \alpha_1, \dots, \alpha_n$ are linearly independent over the

field of rational numbers, then for every $\varepsilon > 0$ there are only finitely many positive integers q with

$$\|q\alpha_1\| \dots \|q\alpha_n\| q^{1+\varepsilon} < 1,$$

where $\|\xi\|$ denotes the distance from a real number ξ to the nearest integer; in particular, we have

$$|\alpha_i - p_i/q| < q^{-(n+1)/n-\varepsilon}, \quad i = 1, \dots, n,$$

for only finitely many n -tuples of rationals $p_1/q, \dots, p_n/q$. A dual to this result is as follows. Let $\alpha_1, \dots, \alpha_n, \varepsilon$ be as before. Then there are only finitely many n -tuples of nonzero integers q_1, \dots, q_n with

$$\|q_1\alpha_1 + \dots + q_n\alpha_n\| \cdot |q_1 \dots q_n|^{1+\varepsilon} < 1.$$

This last theorem can be used to prove that if α is an algebraic number, k a positive integer, and $\varepsilon > 0$, then there are only finitely many algebraic numbers ω of degree at most k such that $|\alpha - \omega| < H(\omega)^{-k-1-\varepsilon}$, where $H(\omega)$ denotes the height of ω . See also [16, 22].

The work of Thue, Siegel, and Roth had the basic limitation of noneffectiveness. A. Baker (1968) succeeded in proving that for any algebraic number θ of degree $n \geq 3$ and any $\kappa > n$, there exists an effectively computable number $c = c(\theta, \kappa) > 0$ such that $|\theta - y/x| > cx^{-n} \exp(\log x)^{1/\kappa}$ for all integers x, y ($x > 0$) [15]. This result is an immediate consequence of the following effective version of a classical theorem on binary Diophantine equations (Thue, 1909): Let $f = f(x, y)$ be an irreducible binary form of degree $n \geq 3$ with integer coefficients, and suppose that $\kappa > n$. Then for any positive integer m , all integer solutions x, y of the equation $f(x, y) = m$ satisfy $\max(|x|, |y|) < c \exp(\log m)^\kappa$, where $c > 0$ is an effectively computable number depending on n, κ , and the coefficients of f . Baker obtained this result by making use of his theorems which give effective estimates of moduli of linear forms in the logarithms of algebraic numbers with algebraic coefficients. A typical theorem reads as follows: Let $\alpha_1, \dots, \alpha_n$ be nonzero algebraic numbers with $\log \alpha_1, \dots, \log \alpha_n$ linearly independent over the rationals, and let β_0, \dots, β_n be algebraic numbers, not all 0, with degrees and heights at most d and H , respectively. Then for any $\kappa > n + 1$, we have $|\beta_0 + \beta_1 \log \alpha_1 + \dots + \beta_n \log \alpha_n| > c \exp(-(\log H)^\kappa)$, where $c > 0$ is an effectively computable number depending only on $n, \kappa, \log \alpha_1, \dots, \log \alpha_n$, and d [14, 19].

Results of this kind have many important applications in number theory. For instance, we obtain a generalization of the Gel'fond-Schneider theorem on transcendental numbers. Furthermore, the imaginary quadratic number fields of class number 1 can be com-

pletely determined on the basis of Baker's result. This was actually done by Baker (1966) and independently by H. M. Stark (1966) (\rightarrow 347 Quadratic Fields).

Refinements and generalizations of Thue's theorem on the finiteness of solutions of binary Diophantine equations have been obtained by Baker and his collaborators. (\rightarrow [17], and also [16]). Also, p - and p -adic analogs of Baker's results are known [18].

There are a number of unsettled problems on the irrationality of particular numbers, such as the Euler constant C , π^e , or $\zeta(2k+1)$ with k a positive integer. R. Apéry (1978) proved that there exist a positive number ε and a sequence of positive integers $\{q_n\}$ such that $0 < \|q_n \zeta(3)\| < e^{-\varepsilon n}$ for all $n \geq 1$, so that $\zeta(3)$ is irrational.

H. Uniform Distribution

Let θ be a real number, x a positive integer, and $[\theta x]$ the maximum integer not larger than θx . We write $(\theta x) = \theta x - [\theta x] \equiv \theta x \pmod{1}$. Jacobi showed that if θ is irrational, then the set $\{(\theta x) | x \in \mathbb{N}\}$ is densely distributed in the interval $(0, 1)$ (\mathbb{N} is the set of positive integers). In general, let f be a real-valued function defined on \mathbb{N} . We say that $f(x) \pmod{1}$ is **uniformly distributed** in the unit interval, or $f(x)$ is **uniformly distributed (mod 1)**, if the following condition is satisfied: Let α, β be an arbitrary pair of real numbers such that $0 \leq \alpha < \beta \leq 1$, and let N be a given positive integer. Let $T(N)$ be the number of positive integers x such that $x \leq N$, $\alpha \leq (f(x)) < \beta$, where $(f(x)) = f(x) - [f(x)]$. Then $\lim_{N \rightarrow \infty} T(N)/N = \beta - \alpha$. In order for $f(x) \pmod{1}$ to be uniformly distributed, it is necessary and sufficient that $\lim_{N \rightarrow \infty} N^{-1} \sum_{x=1}^N e^{2\pi i h f(x)} = 0$ for any non-zero integer h (**Weyl's criterion**, 1914). Weyl proved that if θ is an irrational number, then $\theta x \pmod{1}$ is uniformly distributed.

The following theorem, given by J. G. van der Corput, is often useful: Let f be a real-valued function defined on \mathbb{N} . Consider the function $f_h(x) = f(x+h) - f(x)$ for an arbitrary positive integer h . If $f_h(x) \pmod{1}$ is uniformly distributed (mod 1) for all such h , then $f(x) \pmod{1}$ is also uniformly distributed (mod 1).

Utilizing this theorem, it can be shown that if $f(x) = \theta_r x^r + \theta_{r-1} x^{r-1} + \dots + \theta_1 x$, where at least one of the coefficients θ_i is irrational, then $f(x) \pmod{1}$ is uniformly distributed.

The notion of uniform distribution of sequences of real numbers has an analog in compact Hausdorff spaces and in various topological groups. A systematic treatment of such generalized notions of uniform distribution can be found in [21].

References

- [1] H. Minkowski, *Geometrie der Zahlen*, Teubner, 1910 (Chelsea, 1953).
- [2] H. Minkowski, *Diophantische Approximationen*, Teubner, second edition, 1927 (Chelsea, 1957).
- [3] H. Minkowski, *Gesammelte Abhandlungen I, II*, Teubner, 1911 (Chelsea, 1967).
- [4] J. F. Koksma, *Diophantische Approximationen*, *Erg. Math.*, Springer, 1936 (Chelsea, 1950).
- [5] H. Weyl, C. L. Siegel, and K. Mahler, *Geometry of numbers*, Mimeographed notes, Princeton, 1950.
- [6] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Univ. Press, 1957.
- [7] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer, 1959.
- [8] T. Schneider, *Einführung in die transzendenten Zahlen*, Springer, 1957.
- [9] S. Lang, *Diophantine geometry*, Interscience, 1962.
- [10] C. L. Siegel, *Über einige Anwendungen diophantischer Approximationen*, *Abh. Preuss. Akad. Wiss.*, no. 1 (1929) (*Gesammelte Abhandlungen*, Springer, 1966, vol. 1, 209–266).
- [11] C. L. Siegel, *A mean value theorem in geometry of numbers*, *Ann. Math.*, (2) 46 (1945), 340–347 (*Gesammelte Abhandlungen*, Springer, 1966, vol. 3, 39–46).
- [12] K. F. Roth, *Rational approximations to algebraic numbers*, *Mathematika*, 2 (1955), 1–20.
- [13] A. A. Markoff (A. A. Markov), *Sur les formes quadratiques binaires indéfinies*, *Math. Ann.*, 15 (1879), 381–407.
- [14] A. Baker, *Linear forms in the logarithms of algebraic numbers I, II, III, IV*, *Mathematika*, 13 (1966), 204–216; 14 (1967), 102–107; 14 (1967), 220–228; 15 (1968), 204–216.
- [15] A. Baker, *Contributions to the theory of Diophantine equations I, II*, *Philos. Trans. Roy. Soc. London, A* 263 (1967–1968), 173–191, 193–208.
- [16] A. Baker, *Transcendental number theory*, Cambridge Univ. Press, 1975.
- [17] A. Baker and J. Coates, *Integer points on curves of genus 1*, *Proc. Cambridge Philos. Soc.*, 67 (1970), 595–602.
- [18] A. Brumer, *On the units of algebraic number fields*, *Mathematika*, 14 (1967), 121–124.
- [19] N. I. Fel'dman, *Linear form of logarithmic algebraic numbers (in Russian)*, *Dokl. Akad. Nauk SSSR*, 182 (1968), 1278–1279.
- [20] W. M. Schmidt, *Simultaneous approximation to algebraic numbers by rationals*, *Acta Math.*, 125 (1970), 189–201.

- [21] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Interscience, 1974.
- [22] W. M. Schmidt, *Diophantine approximation*, *Lecture notes in math.* 785, Springer, 1980.
- [23] W. M. Schmidt, *Approximation to algebraic numbers*, *Enseignement Math.*, 17 (1971), 188–253.
- [24] C. F. Osgood (ed.), *Diophantine approximation and its applications*, Academic Press, 1973.

183 (VII.23) Global Analysis

Mathematics that treats, by using functional analytical techniques, various problems concerning the [†]calculus of variations, [†]singularities, infinite-dimensional Lie groups, or [†]non-linear partial differential equations, such as equations of fluids or of gravitation in general relativity, may be called global analysis if it uses as a main tool an infinite-dimensional version of differential geometry and topology analogous to that for finite-dimensional manifolds. The term **global analysis** therefore has no precise definition. However, it can be said that it is analysis on manifolds, and the concept of [†]infinite-dimensional manifolds is the central abstract idea in it.

Suppose one considers a nonlinear differential operator on a finite-dimensional manifold. Then by using functional analytical techniques it often happens that its domain is neither a linear space nor its open subset but an infinite-dimensional manifold, and that such a nonlinear operator can be regarded as a [†]differentiable mapping between infinite-dimensional manifolds. The [†]differential at a point in that source manifold is called a **linearized operator**, to which one can apply various theories of linear functional analysis.

“Global analysis” seems to have first appeared in the literature in the late 1960s [1, 2]. However, the phrase “infinite-dimensional manifold” has been used widely since about 1960. By that early date, local theories of infinite-dimensional manifolds, sometimes called “general analysis,” such as definitions of differentiability, the [†]implicit function theorem, [†]Taylor’s theorem, the existence and uniqueness of [†]ordinary differential equations, and the [†]Frobenius theorem, had already been established in [†]Banach spaces [3]. Therefore, with these regarded as a local theory, the concept of infinite-dimensional manifolds could be defined naturally, and the concept of infinite-

dimensional Lie groups as well [4]. A few years later, R. Palais and S. Smale [5] and J. Eells and J. Sampson [6] showed that such concepts are useful in the calculus of variations, and R. Abraham and J. Robbin [7] remarked that \dagger transversality theorems initiated by R. Thom can be easily proved by using an infinite-dimensional version of \dagger Sard's theorem [8].

The so-called \dagger Atiyah-Singer index theorem [9], announced in 1963, gave impetus to the field as people sought the theorem's most natural expression; the work finally resulted in the theorem classifying separable \dagger Hilbert manifolds by homotopy type (\rightarrow 105 Differentiable Manifolds).

After the appearance of these theorems announcing the nonexistence of differential topology on separable infinite-dimensional Hilbert manifolds, global analysis moved toward more concrete problems and applications to various branches of mathematics. However, many of these applications are formulated not on \dagger Banach or Hilbert manifolds, but on manifolds modeled on Fréchet spaces (\dagger Fréchet manifolds) or on \dagger nuclear spaces. For such situations we have in general no local theories. Neither the implicit function theorem nor the Frobenius theorem holds on such manifolds. However, since these theorems are crucial for nonlinear problems, various kinds of sufficient conditions for the validity of these theorems are being studied by many people (\rightarrow 286 Nonlinear Functional Analysis).

As for the calculus of variations, **Yamabe's problem** is being studied extensively, for this seems to be a typical problem not satisfying the so-called \dagger Condition C. The original paper of H. Yamabe [13], insisting that every compact Riemannian manifold can be conformally deformed into a manifold of constant scalar curvature, contains a serious gap, and the problem is still open, though many cases are known where the statement holds (\rightarrow 364 Riemannian Manifolds H). The harmonic mappings defined in [6] are also being studied extensively (\rightarrow 195 Harmonic Mappings).

In differential geometry or the \dagger general theory of relativity definitions of various curvatures, such as Riemannian, Ricci or scalar, or Gauss, can be sometimes regarded as nonlinear differential equations on manifolds. Proof of the global existence of solutions to these equations has long been sought, and several theorems have recently appeared [14–16].

Infinite-dimensional groups such as $GL(E)$, or $GL_c(E)$ with uniform topology, are called \dagger Banach-Lie groups, and they have been studied in the operator calculus. On the other

hand, infinite groups studied by S. Lie and E. Cartan were in fact infinite-dimensional germs of transformation groups defined on a neighborhood of a point in a manifold. These were not groups in the strict sense. Recently, H. Omori [17] has given a definition of abstract infinite-dimensional Lie groups that includes Banach-Lie groups and many infinite-dimensional transformation groups studied by Cartan. An application of these groups to \dagger fluid dynamics can be found in [18]. Moreover, \dagger unitary representation theories of these groups are now being constructed [19, 20].

Though global analysis consists at present of a rather disorganized combination of many nonlinear problems in analysis on manifolds and in mathematical physics, it is nevertheless one of the most active branches of mathematics.

References

- [1] J. Eells, A setting for global analysis, *Bull. Amer. Math. Soc.*, 72 (1966) 751–807.
- [2] R. Palais, *Foundations of global nonlinear analysis*, Benjamin, 1968.
- [3] J. Dieudonné, *Foundations of modern analysis*, Academic Press, 1960.
- [4] G. Birkhoff, Analytic groups, *Trans. Amer. Math. Soc.*, 43 (1938), 61–101.
- [5] R. Palais and S. Smale, A generalized Morse theory, *Bull. Amer. Math. Soc.*, 70 (1964), 165–172.
- [6] J. Eells and J. Sampson, Harmonic mappings of Riemannian manifolds, *Amer. J. Math.*, 86 (1964), 109–160.
- [7] R. Abraham and J. Robbin, *Transversal mappings and flows*, Benjamin, 1967.
- [8] S. Smale, An infinite-dimensional version of Sard's theorem, *Amer. J. Math.*, 87 (1965), 861–866.
- [9] M. Atiyah and I. Singer, The index of elliptic operators on compact manifolds, *Bull. Amer. Math. Soc.*, 69 (1963), 422–433.
- [10] N. H. Kuiper, The homotopy type of the unitary group of Hilbert space, *Topology*, 3 (1965), 19–30.
- [11] D. Burghlelea and N. H. Kuiper, Hilbert manifolds, *Ann. Math.*, (2) 90 (1969), 335–352.
- [12] J. Eells and D. Elworthy, Open embeddings of certain Banach manifolds, *Ann. Math.*, (2) 91 (1970), 465–485.
- [13] H. Yamabe, On deformation of Riemannian structures on compact manifolds, *Osaka Math. J.*, 12 (1960), 21–37.
- [14] T. Aubin, Equations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire, *J. Math. Pures Appl.*, 55 (1976), 269–296.

- [15] H. Gluck, The generalized Minkowski problem in differential geometry in the large, *Ann. Math.*, 96 (1972), 245–276.
- [16] S. Yau, On the Ricci curvature of a compact Kähler manifold and the complex Monge-Ampère equation I, *Comm. Pure Appl. Math.*, 31 (1978), 339–411.
- [17] H. Omori, Infinite-dimensional Lie transformation groups, *Lecture notes in math.* 427, Springer, 1974.
- [18] D. Ebin and J. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. Math.*, (2) 94 (1970), 101–162.
- [19] A. M. Vershik, I. M. Gel'fand, and M. I. Graev, Representations of the group of diffeomorphisms, *Russian Math. Surveys*, 30 (6) (1975), 1–50. (Original in Russian, 1975.)
- [20] R. S. Ismagilov, Unitary representations of the group of diffeomorphisms of the space \mathbf{R}^n , $n \geq 2$, *Functional. Anal. Appl.*, 9 (1975), 144–155. (Original in Russian, 1975.)

184 (XXI.27) Gödel, Kurt

Kurt Gödel (April 28, 1906–January 14, 1978) was born in Brno, Czechoslovakia (at that time Brünn, Austria-Hungary). He studied mathematics and physics at the University of Vienna, where he took the Ph.D. degree in 1930. After he had taught mathematics at the University of Vienna from 1933 to 1938, he was invited to the Institute for Advanced Study at Princeton, where he became professor in 1953. In 1976, he was named professor emeritus; he died in Princeton in 1978.

Gödel contributed important fundamental results covering all aspects of mathematical logic. Among his famous works are the proof of the \dagger completeness of the first-order predicate calculus, the incompleteness of the consistent axiomatic system containing Peano's arithmetic (incompleteness theorem), and the \dagger consistency of the axiom of choice and the generalized continuum hypothesis. He also introduced the notion of \dagger recursive functions and found the Gödel solution of Einstein's equations of relativity. In addition to mathematical works, he left philosophical papers on set theory and the foundations of mathematics.

Reference

- [1] K. Gödel, The consistency of the axiom of choice and of the generalized continuum hypothesis, Princeton Univ. Press, 1940.

185 (I.8) Gödel Numbers

A. General Remarks

K. Gödel [1] devised the following method to prove his incompleteness theorems (\rightarrow Section C).

Let \mathfrak{S} be a \dagger formal system. In this article, we call its basic symbols, \dagger terms, \dagger formulas, and formal proofs the “constituents” of \mathfrak{S} . Let g be an \dagger injection from the constituents of \mathfrak{S} into the natural numbers satisfying the following two conditions: (1) Given a constituent C , we can compute the value $g(C)$ in a finite number of steps. (2) Given a natural number n , there exists a finitary procedure to find out whether there exists a constituent C of \mathfrak{S} such that $g(C) = n$; furthermore, when such a C exists, it can actually be specified in a finite number of steps.

If such a mapping g is given for the system \mathfrak{S} , then the mapping g is called a **Gödel numbering** and the number $g(C)$ is called the **Gödel number** of the constituent C (with respect to g).

B. An Example of Gödel Numbers

(1) Let $\alpha_0, \alpha_1, \dots$ be the basic symbols of \mathfrak{S} . With each α_i we associate a distinct odd number q_i : $g(\alpha_i) = q_i$ ($i = 0, 1, \dots$). (2) Let F be a constituent of \mathfrak{S} . If F is constructed from any other constituents F_0, F_1, \dots, F_k of \mathfrak{S} by a rule peculiar to \mathfrak{S} (for convenience we write this $F = (F_0, F_1, \dots, F_k)$), and if, for each F_i , $g(F_i)$ is already defined, then we put $g(F) = \langle g(F_0), g(F_1), \dots, g(F_k) \rangle$, where $\langle a_0, a_1, \dots, a_k \rangle$ denotes the number $p_0^{a_0} p_1^{a_1} \dots p_k^{a_k}$ (p_i is the $(i+1)$ st prime number). For example, suppose that \mathfrak{S} contains $0, =, v_j$ (variables), and \neg (negation) among the basic symbols, and let their Gödel numbers be 7, 9, 11^{j+1} , and 13, respectively. Since the formula $\neg(0 = v_j)$ can be analyzed in the form $(\neg, (0, =, v_j))$, its Gödel number is $\langle 13, \langle 7, 9, 11^{j+1} \rangle \rangle$. For details \rightarrow [1, 4].

C. Gödel's Incompleteness Theorems

By means of a Gödel numbering any metamathematical notion about the constituents of a formal system \mathfrak{S} can be interpreted as a number-theoretic notion. For example, the notion “formula” is interpreted as the number-theoretic predicate $\text{Form}(x)$ which means that x is the Gödel number of a formula. The provability of a formula is interpreted as the number-theoretic predicate $\text{Prov}(x)$, which means that x is the Gödel number of a prov-

able formula; accordingly, for any formula A of \mathfrak{S} , the proposition $\text{Prov}(g(A))$ means that A is provable. This interpretation is called the **arithmetization** of metamathematics.

Let the formal system \mathfrak{S} include formal elementary number theory. Then many metamathematically useful number-theoretic predicates can be expressed by the respective formulas of \mathfrak{S} ; for example, there exist formulas **Form**(x) and **Prov**(x) expressing the predicates **Form**(x) and **Prov**(x), respectively. Furthermore, Gödel proved the existence of a closed formula U such that the formula $\neg \text{Prov}(g(U)) \leftrightarrow U$ is provable. This closed formula U is one of the so-called **formally undecidable propositions**, and in fact it is shown that neither U nor $\neg U$ is provable in \mathfrak{S} if \mathfrak{S} is consistent in a strong sense. This result is called Gödel's **first incompleteness theorem**.

By use of the formulas **Form**(x) and **Prov**(x) the consistency of the formal system \mathfrak{S} is expressed by the formula **Consis** which is an abbreviation of $\exists x(\text{Form}(x) \wedge \neg \text{Prov}(x))$. Gödel obtained the result that the formula **Consis** is not provable in \mathfrak{S} if \mathfrak{S} is consistent, on the basis of the following three facts: (1) U is not provable in \mathfrak{S} if \mathfrak{S} is consistent; (2) **Consis** $\rightarrow \neg \text{Prov}(g(U))$ is provable in \mathfrak{S} ; (3) $\neg \text{Prov}(g(U)) \rightarrow U$ is not provable in \mathfrak{S} . This result is called Gödel's **second incompleteness theorem**.

The method of arithmetization is important and useful in the study of mathematical logic. The notion of Gödel numbers of recursive functions is one of its applications (\rightarrow 356 Recursive Functions).

D. Tarski's Theorem Concerning Truth Definitions

Let a consistent formal system \mathfrak{S} and a \dagger model of \mathfrak{S} be given. By means of a Gödel numbering, the truth notion of closed formula is interpreted as the number-theoretic predicate **Tr**(x) which means that x is the Gödel number of a true closed formula; accordingly, for any closed formula A of \mathfrak{S} , the proposition $\text{Tr}(g(A))$ means that A is true.

In relation to the foregoing fact, if there exists a formula **Tr**(x) of a single variable and the formula $\text{Tr}(g(A)) \leftrightarrow A$ is provable for every closed formula A , then that formula **Tr**(x) is called a **truth definition**. A. Tarski proved the fact that there is no truth definition in \mathfrak{S} if the formal system \mathfrak{S} is consistent.

References

[1] K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und ver-

wandter Systeme I, Monatsh. Math. Phys., 38 (1931), 173–198; English translation, On formally undecidable propositions of Principia mathematica and related systems, Oliver & Boyd, 1962.

[2] A. Tarski, Der Wahrheitsbegriff in den formalisierten Sprachen, *Studia Philosophica*, 1 (1936), 261–405, translated from the Polish original 1933; English translation, The concept of truth in formalized languages, logic, semantics, metamathematics, Oxford Univ. Press, 1956.

[3] J. B. Rosser, Extensions of some theorems of Gödel and Church, *J. Symbolic Logic*, 1 (1936), 87–91.

[4] S. C. Kleene, Introduction to metamathematics, North-Holland and Noordhoff, 1952.

[5] S. C. Kleene, Mathematical logic, Wiley, 1967.

[6] J. R. Shoenfield, Mathematical logic, Addison-Wesley, 1967.

186 (XVI.12) Graph Theory

A. Overview of Graph Theory

Two aspects of graphs are the object of **graph theory**. One is that a graph expresses a binary relation over a set V , and the other is the fact that a graph is a CW-complex of 1 dimension, an object of study in algebraic topology. Because of its special nature as an object of 1 dimension, we can consider various concrete properties in detail. Hence graph theory has close connection with other areas, such as network theory, system theory, automata theory, and the theory of computational processes, and it has many useful applications.

The notion of a “graph” as currently used in graph theory was first discussed by L. Euler [1]. It is said that J. J. Sylvester coined the word “graph” as we presently understand it. However, up to now, there were few unifying principles, and graph theory seemed like a large collection of miscellaneous problems and ad hoc techniques. The terminology has not yet been standardized; different usages prevail in different schools. The resulting confusion can be seen in reference books such as [2]–[5]. Recently, infinite graphs have also been studied. But here we restrict ourselves to finite graphs, since only they give a typical theory.

B. Definition of Graph

The notion of a **graph** $G = (V, E, \delta^+, \delta^-)$ is a composite notion of two finite sets V and E

and two maps $\partial^+ : E \rightarrow V$ and $\partial^- : E \rightarrow V$. An element of V is called a **vertex**, and an element of E is called an **edge**. The maps ∂^\pm are called **incidence relations**. The terms **point**, or **node** instead of vertex, and **arc**, **line**, or **branch** instead of edge have also been used (sometimes with slightly different meanings). For an edge $e \in E$, $\partial^+ e \in V$ is called the **initial vertex** and $\partial^- e \in V$ is called the **terminal vertex**. Both are called the **end vertices** of e . The inverse map δ^\pm of $\partial^\pm : V \rightarrow 2^E$ is defined by $\delta^\pm v = \{e \in E | \partial^\pm e = v\}$ and has the following properties: (i) If $v \neq v'$, then $\delta^+ v \cap \delta^+ v' = \delta^- v \cap \delta^- v' = \emptyset$. (ii) $\bigcup_{v \in V} \delta^+ v = \bigcup_{v \in V} \delta^- v = E$. Conversely, if the maps $\delta^\pm : V \rightarrow 2^E$ have the properties (i) and (ii), then there exist corresponding maps ∂^\pm . For a vertex $v \in V$, $|\delta^+ v|$ is called the **outdegree** or **positive degree**, $|\delta^- v|$ is called the **indegree** or **negative degree**, and the sum $|\delta^+ v| + |\delta^- v|$ is called the **degree**. We always have the relation $\sum_{v \in V} |\delta^+ v| = \sum_{v \in V} |\delta^- v| = |E|$, and $\sum_{v \in V} (|\delta^+ v| + |\delta^- v|) = 2|E|$. The number of vertices with odd degree is always even. Two end vertices of an edge are called mutually **adjacent**, and two edges with at least one common end vertex are also called mutually **adjacent**. An edge satisfying $\partial^+ e = \partial^- e$ is called a **self-loop**, and a vertex satisfying $\delta^+ v = \delta^- v = \emptyset$ is called an **isolated vertex**.

When permutation groups P_V operating over V and P_E operating over E are given, we can naturally define the permutation (π_V, π_E) over the graph $G(\pi_V \in P_V, \pi_E \in P_E)$. A graph is classified into equivalence classes by P_V, P_E . When P_V, P_E are all the permutations of V, E , respectively, each equivalence class is called an **unlabeled graph**, whereas when P_V, P_E consist only of an identity transformation, G is called a **labeled graph**.

For a given graph $G = (V, E, \delta^+, \delta^-)$ and a subset $E' \subset E$, we can define the **reoriented graph** of edges in E' by $G' = (V, E, \delta'^+, \delta'^-)$, where $\delta'^+ e = \partial^+ e$ if $e \notin E'$ and $\delta'^\pm e = \partial^\mp e$ if $e \in E'$. We have an equivalence relation by identifying the reoriented graphs. Each class by this equivalence relation is called an **undirected graph** or an **unoriented graph**. In contrast, the graph in the original sense is called a **directed graph** or an **oriented graph**.

C. Examples of Special Graphs

(1) A **complete graph** is a graph such that there exists one and only one edge with end vertices v and v' for every different pair of vertices v and v' . (2) A **bipartite graph** is a graph with a partition $V^+ \cup V^- = V$ such that $V^+ \cap V^- = \emptyset$ and for every $e \in E$, we have $\partial^+ e \in V^+$ and $\partial^- e \in V^-$. If there always exists an edge e with $\partial^+ e = v, \partial^- e = v'$ for every pair $v \in V^+$ and

$v' \in V^-$, it is called a **complete bipartite graph**. (3) A **regular graph** is a graph whose degree at each vertex is the same. (4) A **partial graph** is defined as follows: Let E' be a subset of E in a graph $G = (V, E, \delta^+, \delta^-)$. The graph $G' = (V', E', \delta'^+, \delta'^-)$ is called a **partial graph** of G , where $V' = \partial^+ E' \cup \partial^- E'$, and δ'^\pm are the restrictions of δ^\pm on E' , respectively. When $E' = E$, the partial graph is the graph obtained by deleting all isolated vertices from G . Similarly, for a subset V' of V , the graph $G'' = (V'', E'', \delta''^+, \delta''^-)$, defined by $E'' = \delta^+ V'' \cup \delta^- V''$ and with δ''^\pm being the restriction of δ^\pm on E'' , is called a **section graph** of G .

D. Representation of a Graph

When we want to process by computer a problem concerning a graph, we must represent a graph in a suitable form (\rightarrow 96 Data Processing). A commonly used method is the following **list representation**: Let $V = \{v_1, v_2, \dots, v_M\}$ and $E = \{e_1, e_2, \dots, e_n\}$. For each $v_\alpha \in V$ ($\alpha = 1, \dots, M$), we arrange the edges in $\delta^+ v_\alpha$ and in $\delta^- v_\alpha$ in suitable orders. Then (i) for each $v_\alpha \in V$, record $|\delta^+ v_\alpha|, |\delta^- v_\alpha|, B_\alpha^+, \bar{B}_\alpha^+, B_\alpha^-, \bar{B}_\alpha^-$, where B_α^\pm and \bar{B}_α^\pm are the first and the last index of the edges in $\delta^\pm v_\alpha$, respectively. We define the corresponding values to be 0 if the $\delta^\pm v_\alpha$ are empty. (ii) For each $e_\kappa \in E$, record $\partial_\kappa^+, \partial_\kappa^-, E_\kappa^+, \bar{E}_\kappa^+, E_\kappa^-, \bar{E}_\kappa^-$, where ∂_κ^\pm are the number of vertices $\partial^\pm e_\kappa$, and $E_\kappa^\pm, \bar{E}_\kappa^\pm$ are the number of edges immediately after or before e_κ among the edges with the same initial (for $^+$) or terminal (for $^-$) vertex as that of e_κ . If e_κ is the last or the first among such edges, we define the corresponding values to be 0.

E. Deformation of a Graph

Let $G = (V, E, \delta^+, \delta^-)$ be a graph. A graph obtained by **opening** an edge e is the graph $G' = (V, E - \{e\}, \delta'^+, \delta'^-)$, where δ'^\pm is the restriction of δ^\pm on $E - \{e\}$. The graph obtained by deleting all isolated vertices from G' is a partial graph of G over $E - \{e\}$. A graph obtained by **shortening** an edge e is the graph $G'' = (V'', E - \{e\}, \delta''^+, \delta''^-)$, where $V'' = (V - \{\partial^+ e, \partial^- e\}) \cup \{\delta\}$, δ being a new vertex not contained in V , and $\delta''^\pm = \varphi \cdot \delta^\pm$, $\varphi : V \rightarrow V''$ being defined by $\varphi(v) = v$ for $v \neq \partial^\pm e_\kappa$ and $= \delta$ when $v = \partial^+ e_\kappa$ or $v = \partial^- e_\kappa$. A graph obtained by opening and shortening several different edges is similarly defined, and the result is independent of the order of opening or shortening process. A graph obtained by opening edge(s) or by shortening edge(s) or by both processes is called a **subgraph**, a **contraction**, or a **subcontraction**, respectively.

F. Connectedness

Let $G=(V,E,\partial^+,\partial^-)$ be a graph, and let v_α and v_β be two vertices. A **path** from v_α to v_β of **length** l is a sequence $P=(v_\alpha=v_{\alpha_0},e_1e_{\kappa_1},v_{\alpha_1},e_2e_{\kappa_2},\dots,e_l e_{\kappa_l},v_{\alpha_l}=v_\beta)$, where e_i are $+1$ or -1 , and for every $i=1,\dots,l$, we have $\partial^+e_{\kappa_i}=v_{\alpha_{i-1}}$, $\partial^-e_{\kappa_i}=v_{\alpha_i}$ if $e_i=+1$, and $\partial^+e_{\kappa_i}=v_{\alpha_i}$, $\partial^-e_{\kappa_i}=v_{\alpha_{i-1}}$ if $e_i=-1$. When $v_\alpha=v_\beta$, it is called a **closed path** considering P as a cyclic sequence. If in the sequence P of a path no edge appears more than once, it is called a **simple path**. Similarly, if no vertex appears more than once, it is called **elementary**. If all $e_i=+1$, it is called a **direct path**. **Direct closed paths**, etc., are similarly defined.

If we define $v\sim v'$ by the existence of a path from v to v' , this \sim is an equivalence relation. Let us denote the equivalence classes by V_1,\dots,V_s . The section graph $G_i=(V_i,E_i,\partial_i^+,\partial_i^-)$ determined by V_i is called a **connected component**, or simply a **component**, of G . The sets E_1,\dots,E_s are mutually disjoint and the union is E . If we denote by $v\rightarrow v'$ the existence of a direct path from v to v' , the relation \rightarrow is a \dagger pseudo-order. The relation $v\leftrightarrow v'$ defined by $v\rightarrow v'$ and $v'\rightarrow v$ is an equivalence relation in V , and the equivalence classes $\tilde{V}_1,\dots,\tilde{V}_t$ or the section graphs $\tilde{G}_i=(\tilde{V}_i,\tilde{E}_i,\tilde{\partial}_i^+,\tilde{\partial}_i^-)$ determined by \tilde{V}_i are called **strongly connected components** of G . $\tilde{E}_1,\dots,\tilde{E}_t$ are mutually disjoint, but their union is not always E . Among the sets $\tilde{V}_1,\dots,\tilde{V}_t$, we can define an order relation $\tilde{V}_i\rightarrow\tilde{V}_j$ by the existence of $v\in\tilde{V}_i, v'\in\tilde{V}_j, v\rightarrow v'$. Classification by \leftrightarrow is a refinement of that by \sim . When $s=1$, the graph G is called **connected**, and when $t=1$, G is called **strongly connected**.

For a different pair $v, v'\in V$, a set $S\subset V-\{v,v'\}$ is called a **separator** of v and v' if every path from v to v' contains at least one vertex of S . When every separator S for any pair v, v' has at least k elements, the graph G is called **k -connected**. For $k\geq 3$, there are several variations of the definition of k -connectedness. 1-connectedness is equivalent to connectedness in the sense defined above.

A simple path containing all edges of a graph is called an **Euler path**. Euler direct paths, etc., are similarly defined. A graph with an Euler path is called an **Euler graph**. A graph is an Euler graph if and only if (i) it is connected, and (ii) the number of vertices with odd degrees is 0 or 2 (Euler's unicursal graph theorem [1]). Similar results are known for Euler closed paths or Euler direct paths. An elementary path containing all vertices of a graph is called a **Hamilton path**. Hamilton closed paths, etc., are similarly defined. The criteria for the existence of a Hamilton path for any given graph is unknown. This is known to be an \dagger NP-complete problem (\rightarrow 71

Complexity of Computations). Various sufficient conditions or necessary and sufficient conditions for special graphs have been given (\rightarrow e.g., [6]).

G. Tieset, Cutset, Tree, and Cotree

For a graph $G=(V,E,\partial^+,\partial^-)$, we define its **incidence matrix** $[\bar{D}_\kappa^\alpha(\alpha=1,\dots,M(=|V|),\kappa=1,\dots,n(=|E|))]$ by defining D_κ^α to be 1 if $v_\alpha=\partial^+e_\kappa\neq\partial^-e_\kappa$, -1 if $v_\alpha=\partial^-e_\kappa\neq\partial^+e_\kappa$, and 0 if $\partial^+e_\kappa=\partial^-e_\kappa$ or $e_\kappa\notin\partial^+v_\alpha\cup\partial^-v_\alpha$. Similarly we define its **adjacency matrix** $[\Gamma_\beta^\alpha(\alpha,\beta=1,\dots,M)]$ by defining Γ_β^α to be 0 if v_α and v_β are not mutually adjacent and 1 if v_α and v_β are mutually adjacent.

A set of edges forming a closed path is called a **tieset**. A set of edges of the form $\{e|\partial^+e\in W, \partial^-e\in V-W\}\cup\{e|\partial^-e\in W, \partial^+e\in V-W\}$ for a partition $(W,V-W)$ of V is called a **cutset**. A maximal subset of edges containing no tiesets is called a **tree**, and a maximal subset of edges containing no cutsets is called a **cotree**. A tree is sometimes called a **spanning tree** of G . Every tree is the complement (with respect to E) of a cotree, and vice versa.

The number of elements of a tree is always the same and equal to the \dagger rank m of the incidence matrix. This value m is called the **rank** of the graph G . Similarly, the number k of elements of a cotree is always the same, which is called the **nullity** or the **cyclomatic number** of the graph G . Always, $k=n-m$.

Let K be a field, K^n, K^M be vector spaces over K of dimensions n and M , respectively, and K^{n*}, K^{M*} be the \dagger dual spaces of K^n, K^M . The incidence matrix defines two mutually contragradient linear mappings $\partial:K^n\rightarrow K^M$ and $\delta:K^{M*}\rightarrow K^{n*}$ with respect to their canonical bases. A minimal set among the family of supports (in E) of nonzero vectors in the kernel $\text{Ker } \partial$ is a tieset corresponding to an elementary closed path. Similarly, a minimal set among the family of supports (in E) of nonzero vectors of the image $\text{Im } \delta$ is a minimal cutset.

Let $T(\subset E)$ be a tree and $\bar{T}=E-T$ be a cotree. Let us renumber the edges so that $T=\{e_1,\dots,e_m\}$, $\bar{T}=\{e_{m+1},\dots,e_n\}$. For each $e_{m+p}\in\bar{T}$, there is a unique vector $R_p^k(p=1,\dots,k(n-m))$ in $\text{Ker } \partial$ whose support is in $\{e_{m+p}\}\cup T$ and whose e_{m+p} - equals $+1$. Similarly, for each $e_a\in T(a=1,\dots,m)$, there is a unique vector $D_a^k(a=1,\dots,m)$ in $\text{Im } \delta$ whose support is in $\{e_a\}\cup\bar{T}$ and whose e_a - component equals $+1$. $\{R_1^k,\dots,R_k^k\}$ is a basis of $\text{Ker } \partial$, and $\{D_1^k,\dots,D_m^k\}$ is a basis of $\text{Im } \delta$. Both matrices $R_p^k(k\times n)$ and $D_a^k(m\times n)$ are **totally unimodular**, i.e., every minor determinant is 0, $+1$, or -1 . Furthermore, the following relations hold: $R_p^{m+q}=\delta_p^q(p,q=1,\dots,k), D_b^a=\delta_b^a(a,b=1,\dots,m), R_p^a+$

$D_{m+p}^a = 0$ ($a = 1, \dots, m; p = 1, \dots, k$). The matrices R_p^k and D_k^a are called the **fundamental tieset matrix** and the **fundamental cutset matrix**, respectively, with respect to the tree-cotree pair (T, \bar{T}) . The minor of the matrix R_p^k consisting of all rows and k columns $\kappa_1, \dots, \kappa_k$ is $+1$ or -1 if and only if $\{e_{\kappa_1}, \dots, e_{\kappa_k}\}$ is a cotree, and 0 otherwise. Similarly, the minor of the matrix D_k^a consisting of all rows and m columns $\kappa_1, \dots, \kappa_m$ is $+1$ or -1 if and only if $\{e_{\kappa_1}, \dots, e_{\kappa_m}\}$ is a tree, and 0 otherwise.

H. Planarity of a Graph

Let there be a natural one-to-one correspondence between the sets of edges of two graphs $G_i = (V_i, E_i, \partial_i^+, \partial_i^-)$ ($i = 1, 2$), and suppose that under the correspondence a tree T_1 in G_1 corresponds to a tree T_2 in G_2 and that the fundamental cutset matrices D_1 and D_2 with respect to $(T_i, E_i - T_i)$ are mutually equal. Then G_1 and G_2 are said to be **2-isomorphic**. The definition is equivalent to the one given by the equality of fundamental tieset matrices. 2-isomorphism is an equivalence relation. If G_1 and G_2 are 2-isomorphic, the families of trees, cotrees, tiesets and cutsets are mutually corresponding. The coincidence of one of the families is a sufficient condition for the 2-isomorphism of G_1 and G_2 as undirected graphs. A 3-connected graph has no 2-isomorphic graph other than itself.

Similarly, when a tree T_1 of G_1 corresponds to a cotree \bar{T}_2 of G_2 and if the fundamental cutset matrix of G_1 with respect to $(T_1, E_1 - T_1)$ is equal to the fundamental tieset matrix of G_2 with respect to $(E_2 - T_2, T_2)$ as matrices, then we say that G_2 is **dual** to G_1 . In this case, G_1 is dual to G_2 also. If G_1 is dual to G_2 and G_2 is dual to G_3 , then G_1 and G_3 are 2-isomorphic. The duality is a relation among the equivalence classes of graphs by 2-isomorphisms.

Any graph $G = (V, E, \partial^+, \partial^-)$ can be "drawn" in 3-dimensional Euclidean space in the following sense: Each vertex is a (distinct) point, and an edge e is an arc connecting two points $\partial^+ e$ and $\partial^- e$ with a direction pointing from $\partial^+ e$ to $\partial^- e$ in such a way that no two arcs intersect. A graph representable on a plane or on a 2-dimensional sphere in the above sense is called a **planar graph**. A graph G is planar if and only if it has a dual graph (H. Whitney [7]). Another necessary and sufficient condition is that as an undirected graph, neither the complete five-point graph K_5 nor the bipartite complete graph of three-three points $K_{3,3}$, appear in any subcontraction of the graph G . (This is a version of the criterion of C. Kuratowski [8].)

I. Coloring of Graphs

A **coloring** of the vertices of a graph $G = (V, E, \partial^+, \partial^-)$ is a mapping ψ from V to the set of integers N satisfying the condition $\psi(v) \neq \psi(v')$ for all adjacent vertices v and v' . $\gamma(G) = \min\{|\psi(V)| \mid \psi \text{ is a coloring of the vertices of } G\}$ is called the **chromatic number** of G . If the graph G is drawn over a 2-dimensional closed surface with a 1-dimensional \dagger Betti number b , we have $\gamma(G) \leq \lfloor (7 + \sqrt{1 + 24b})/2 \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integral part. Except for the \dagger Klein bottle (with $b = 2$), where $\gamma(G) \leq 6$, this is the best possible, i.e., there exists a graph whose $\gamma(G)$ equals the upper bound on the right-hand side. As for $b \geq 1$, the inequality was first proved by P. J. Heawood (1890), and final results were established by J. W. T. Young and G. Ringel [10]. When $b = 0$, $\gamma(G) \leq 5$ was shown by A. B. Kempe (1879), but the **four color conjecture**: " $\gamma(G) \leq 4$?" has remained unsolved for more than a hundred years. The conjecture is believed to have been solved affirmatively recently through checking a huge number of cases on a large computer [11, 12].

A subset $W \subset V$ is called an **independent set** or an **internally stable set** if no two vertices in W are mutually adjacent. $\alpha(G) = \max\{|W| \mid W \text{ is an independent set of } G\}$ is called the **number of independence** of G . A subset W of V is called a **dominating set** or **externally stable set** of G if every vertex $v \in V$ is either $v \in W$ or adjacent to a vertex of W . $\beta(G) = \min\{|W| \mid W \text{ being a dominating set of } G\}$ is called the **number of domination** of G . For every graph G , we have $\alpha(G) \geq \beta(G)$ and $\gamma(G) \cdot \alpha(G) \geq |V|$.

J. Decision Problems and Graphs

There are many interesting topics in decision problems concerning graphs, especially from the standpoint of \dagger complexity of computations (\rightarrow e.g., [13]). The following are some typical problems: (i) Problems for which algorithms of polynomial order are known: Is the given graph k -connected, strongly connected, a Euler graph, or a planar graph? (ii) \dagger NP-complete problems: Is the given graph a Hamilton graph? Do we have $\alpha(G) = k$, $\beta(G) = k$, or $\gamma(G) = k$?

Let G_1 and G_2 be two graphs. The problem of whether they coincide as unlabeled graphs is called the **isomorphism problem**, and has been studied for many years in connection with problems concerning the structure of chemical compounds. Unfortunately, no algorithm of polynomial order is known; nor do we know whether this is an NP-complete problem. As for the isomorphism problem for

planar graphs, algorithms of polynomial order are known.

K. Perfectness Theorem

Let the number of independence and the chromatic number of a graph $G = (V, E, \partial^+, \partial^-)$ be $\alpha(G)$ and $\gamma(G)$, respectively. Let V_1, \dots, V_r be a disjoint decomposition of V . We denote by $\theta(G)$ the minimal number of r such that every section graph of G over V_i contains a complete graph. Furthermore, we denote by $\omega(G)$ the maximum value of $|W|$ for a subset $W \subset V$ such that the section graph of G over W contains a complete graph. We always have $\alpha(G) \leq \theta(G)$ and $\gamma(G) \leq \omega(G)$. G is called α -**perfect** if every section graph H of G satisfies $\alpha(H) = \theta(H)$. Similarly, G is called γ -**perfect** if $\gamma(H) = \omega(H)$ for every section graph H of G . The conjecture that γ -perfectness and α -perfectness are equivalent has recently been solved affirmatively [14]. This is called the **perfectness theorem**.

References

- [1] L. Euler, *Solutio problematis ad geometriam situs pertinentis*, *Commentarii Academiae Petropolitanae*, 8 (1736), 128–140.
- [2] F. Harary, *Graph theory*, Addison-Wesley, 1969.
- [3] C. Berge, *Graphes et Hypergraphes*, Dunod, second edition, 1974.
- [4] A. A. Zykov, *Finite graph theory* (in Russian), Nauka, 1969.
- [5] M. Iri, *Network flow, transportation and scheduling*, Academic Press, 1969.
- [6] H. Walther and H.-J. Voss, *Über Kreise in Graphen*, VEB Deutscher Verlag der Wissenschaften, 1974.
- [7] H. Whitney, *Non-separable and planar graphs*, *Trans. Amer. Math. Soc.*, 34 (1932), 339–362.
- [8] C. Kuratowski, *Sur le problème des courbes gauches en topologie*, *Fund. Math.*, 15 (1930), 271–283.
- [9] F. Harary and W. Tutte, *A dual form of Kuratowski's theorem*, *Canad. Math. Bull.*, 8 (1965), 17–20, 373.
- [10] G. Ringel, *Map color theorems*, Springer, 1974.
- [11] K. Appel and W. Haken, *Every planar map is four colourable, I. Discharging*, *Illinois J. Math.*, 21 (1977), 429–489; K. Appel, W. Haken and J. Koch, *II. Reducibility*, *Illinois J. Math.*, 21 (1977), 490–567.
- [12] S. Hitotumatu, *Four color problem* (in Japanese), Kodansha, 1978.
- [13] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *The design and analysis of computer algorithms*, Addison-Wesley, 1974.
- [14] L. Lovász, *Normal hypergraphs and perfect graph conjecture*, *Discrete Math.*, 2 (1972), 253–267.

187 (XXI.2) Greek Mathematics

It is generally believed that theoretical mathematics originated with the Greeks. The Greeks learned the arts of land surveying and commercial arithmetic from earlier civilizations; but they developed theoretical mathematics themselves, toward the middle of the 4th century B.C. The creation of a mathematics that transcends practical purposes was one of the most remarkable events in the history of human culture, and one that had an immense impact on the development of all branches of science. We owe the reestablishment of important Greek mathematical texts and the reconstruction of the development of Greek mathematics to the historians of the 19th century (the oldest extensive exposition of the history of mathematics before Euclid is due to Proklos (Proclus) (410–485)).

The earliest known Greek mathematicians are Thales of Miletus (c. 639–546 B.C.) and Pythagoras of Samos (fl. 510? B.C.). Both were Ionians, but the latter went to what is now southern Italy and founded a semireligious school whose members called themselves Pythagoreans. Their motto was “Everything is number”; their studies were called *mathema* (“what is learned”) and consisted of music, astronomy, geometry, and arithmetic (the subject group called the **quadriivium**, which formed the core of medieval and later higher education) “for the purification of soul.” Their research delved into the theories of proportion (in relation to music) and [†]polygonal numbers (triangular numbers, square numbers, etc.), and more generally into the theory of numbers and geometric algebra. It is said that they knew of the irrationality of $\sqrt{2}$, though no evidence of this has been found. Even after the demise of the Pythagorean school, its followers continued to promote mathematics in collaboration with the Academy of Plato.

Another significant school was the Eleatic. Among its members Zeno (c. 490–c. 430 B.C.) is especially important. Zeno's paradoxes are arguments leading to absurdity. Some see within them the origin of logical reasoning and, consequently, of theoretical mathematics [3]. It is chronologically difficult to attribute

to Zeno the consideration of the continuum and irrational numbers, but we can find in him the impetus toward atomistic reasoning. The computation of the volume of pyramids (by dividing them into “atomistic” laminae) by Democritus (fl. 430? B.C.) and the atomistic calculation of the area of circles by Antiphon (fl. 430 B.C.) came shortly after the time of Zeno.

The middle decades of the 4th century B.C. are known as the Age of Pericles, the Golden Age of Athens. The †trisection of an angle, the †duplication of a cube, and the †quadrature of a circle, known at that time as the “**three big problems**” (→ 179 Geometric Construction), were studied by the Sophists. Hippias of Elis (fl. 420 B.C.), Hippocrates of Chios (fl. 430 B.C. in Athens), Archytas of Taras (c. 430–365 B.C.), Menaechmus (fl. 350 B.C.), and his brother Dinostratus (fl. 350 B.C.) solved these problems using conic sections and the **quadratrix** (a transcendental curve whose equation is $y = x \cot(\pi x/2)$).

By 400 B.C. Athens had lost its political influence, but it remained the center of Greek culture. It was during this time that Plato’s Academy flourished, and Plato (427–347 B.C.) and his followers laid particular importance on mathematics. Archytas, Menaechmus, and Dinostratus belonged to or were closely associated with this school. During the first fifty years of the Academy, research in the following fields was pursued: methodology of mathematics or science in general (i.e., dialectics, analysis, synthesis); geometric reconstruction of Mesopotamian algebra; the theory of irrationals in relation to the geometrization of algebra (Theodorus of Cyrene (5th century B.C.), who was Plato’s teacher, as well as Theaitetus of Athens (415?–369 B.C.) contributed to this study, and the general theory of proportion by Eudoxus of Cnidos (c. 408–c. 355 B.C.) also belongs to this field); the method of exhaustion (by Eudoxus); and studies of the “three big problems” and conic sections. It was this school in which the term *mathema* came to be used in its present sense of “mathematics” rather than in the sense of disciplines in general.

The conquests of Alexander the Great accelerated the already considerable cultural influence of Athens. Later, during the Ptolemaic period, the center of culture moved to Alexandria. The Mouseion at Alexandria, the combined library and university, is said to have possessed hundreds of thousands of volumes.

At Alexandria, Euclid (c. 300 B.C.) compiled his *Elements*, which became a model for scientific works for centuries to come—Newton’s *Principia* as well as Spinoza’s *Ethics*

are modeled on it. Most historians say that Euclid’s method derives from Aristotle (384–322 B.C.), who, after studying at Plato’s Academy, founded a new school, the Peripatetics, whose doctrines are at many points opposed to those of the Academy. Others, however, see the origin of Euclid’s axiomatic method in the Eleatics [5]; we can find prototypes of some parts of the *Elements* in both Oenopides, who lived during the time of Zeno, and in Hippocrates.

The third century B.C. was the Golden Age of Greek mathematics. Archimedes of Syracuse (c. 282–212 B.C.) was the greatest mathematician, mechanic, and technician of antiquity. He did important work in mathematics, studying the exact quadrature of the parabola. According to his *epodos* (method), he would obtain a result by mechanical experiments and then prove it by the method of “exhaustion.” He also computed the value of π ; studied spirals and other curves, spheres, and circular cylinders; contributed to the development of statics and optics and their application; and had a profound influence on later mathematicians. During the same period, Apollonius of Perga (fl. 210 B.C.) wrote *Konikon biblia* (Books on Conics) in eight books, of which the last has been lost. The geometric theory of †conic sections contained in this work is not much different from the one we know today; it had a great influence on 17th-century scientists especially. Other mathematicians of this period worth noting are Eratosthenes (c. 275–195 B.C.), who conceived the †sieve method of finding prime numbers and who measured the earth, and Hipparchus (fl. 150 B.C.), called the father of astronomy, who made a table of sines.

Hellenistic influence began to decline in the first century B.C., and the influence of Alexandria decreased. The Mouseion burned in 48 B.C., but was rebuilt. Among the mathematicians of this time, we may count Heron (fl. 60? A.D.); Menelaus (fl. 100 A.D.), who wrote *Sphaerica*; Theon of Smyrna (fl. 125 A.D.); Ptolemy (fl. 150 A.D.), the author of *Almagest*; Nicomachus (50?–150? A.D.), the author of *Arithmetike eisagoge*; Diophantus (fl. 250? A.D.), whose career is not fully known but who wrote *Arithmetika*, of which six of the original thirteen books remained to influence †Fermat; and Pappus (fl. 300 A.D.), the last creative mathematician in Greece, who left eight books of the *Synagoge*, which influenced †Descartes and which still exist today.

The period following the fall of the Western Roman Empire was a difficult one for Greco-Egyptian science. The Mouseion was destroyed for the second time in 392 A.D. Theon of Alexandria (fl. 380) and his daughter, Hy-

patia (c. 370–415), were at that time working on commentaries on the classics. Among the few remaining works of the period is Proclus' (410–485) commentaries on the first book of Euclid's *Elements*. The Athenian Academy was closed in 529 by order of the Emperor Justinian; the last director was Simplicius, who commented on Aristotle. Soon afterward, Alexandria fell into the hands of the Moors, and many scholars fled as refugees to Constantinople, the capital of the Eastern Empire.

References

- [1] T. L. Heath, *A history of Greek mathematics I, II*, Clarendon Press, 1921.
 - [2] T. L. Heath, *A manual of Greek mathematics*, Oxford Univ. Press, 1931 (Dover, 1963).
 - [3] M. Cantor, *Vorlesungen über Geschichte der Mathematik I*, Teubner, third edition, 1907.
 - [4] B. L. van der Waerden, *Zenon und die Grundlagenkrise der griechischen Mathematik*, *Math. Ann.*, 117 (1940), 141–161.
 - [5] Á. Szabó, *Anfänge des euklidischen Axiomensystems*, *Arch. History Exact Sci.*, 1 (1960), 37–106.
 - [6] J. L. Heiberg (ed.), *Euclidis opera omnia I–XIII and suppl.*, Leipzig, 1883–1916.
 - [7] T. L. Heath, *The thirteen books of Euclid's Elements I, II, III*, Cambridge Univ. Press, 1908 (Dover, 1956).
 - [8] P. Ver Eecke (trans.), *Les oeuvres complètes d'Archimède*, Desclée de Brouwer & Cie, 1921 (Blanchard, 1961).
 - [9] P. Ver Eecke (trans.), *Proclus Diadochus, les commentaires sur le premier livre des Éléments d'Euclide*, Desclée de Brouwer & Cie, 1948 (Blanchard, 1959).
 - [10] P. Tannery, *Le géométrie grecque, comment son histoire nous est parvenue et ce que nous en savons*, Paris, 1887.
 - [11] B. L. van der Waerden, *Science awakening*, Noordhoff, 1954.
 - [12] M. Clagett, *Greek science in antiquity*, Abelard-Schuman, 1955 (Collier, 1963).
 - [13] O. Becker (ed.), *Zur Geschichte der griechischen Mathematik*, Darmstadt, 1965.
 - [14] Á. Szabó, *Anfänge der griechischen Mathematik*, Oldenbourg, 1969.
- For the history of mathematics in general,
- [15] R. C. Archibald, *Outline of the history of mathematics*, Mathematical Association of America, sixth edition, 1949.
 - [16] N. Bourbaki, *Eléments d'histoire des mathématiques*, Hermann, second edition, 1969.
 - [17] M. Cantor, *Vorlesungen über Geschichte der Mathematik I–IV*, Teubner, second edition, 1894–1908.

[18] E. Montucla, *Histoire des mathématiques I–IV*, Paris, new edition, 1799–1802.

[19] D. E. Smith, *A source book in mathematics*, McGraw-Hill, 1929 (Dover, 1959).

[20] D. J. Struik, *A concise history of mathematics*, Dover, third edition, 1967.

188 (XIII.28) Green's Functions

A. General Remarks

Green's functions are usually considered in connection with [†]boundary value problems for ordinary differential equations and also with [†]elliptic and [†]parabolic partial differential equations. For example, consider boundary value problems for the [†]Laplacian in 3-dimensional space: $L[u] = (\partial^2/\partial x_1^2 + \partial^2/\partial x_2^2 + \partial^2/\partial x_3^2)u$. Let D be a bounded domain with a smooth boundary S and the boundary condition B on S be either $u(x) = 0$ ($x \in S$) (the first kind) or $\partial u/\partial n + \beta u = 0$ ($x \in S$) (the third kind), where n is the outer normal of unit length, $\beta(x) \geq 0$, and $\beta(x) \neq 0$. We say that the function $g(x_1, x_2, x_3; \xi_1, \xi_2, \xi_3)$ is the **Green's function** of L (or the partial differential equation $L[u] = 0$) relative to the boundary condition B , when (i) $g(x, \xi)$ satisfies $L_x[g(x, \xi)] = 0$ except for $x = \xi$; (ii) $g(x, \xi) = -1/4\pi r + \omega(x, \xi)$, where $r = (\sum_{i=1}^3 (x_i - \xi_i)^2)^{1/2}$ and $\omega(x, \xi)$ is a regular function, i.e., of class C^v for a suitable value v ; (iii) $g(x, \xi)$ satisfies the boundary condition B , i.e., $g(x, \xi) = 0$, $x \in S$ (the first kind), or $(\partial/\partial n + \beta)g(x, \xi) = 0$, $x \in S$ (the third kind). Conditions (i) and (ii) mean that $g(x, \xi)$ is a [†]fundamental solution of L , i.e., $L_x[g(x, \xi)] = \delta(x - \xi)$, where $\delta(x - \xi)$ is [†]Dirac's measure at the point $x = \xi$. Note that if $g(x, \xi)$ is a fundamental solution, then by adding any solution u of the equation $L[u] = 0$ to g , we obtain another fundamental solution $g + u$. Thus Green's function is the fundamental solution that satisfies the given boundary condition. To be more precise, in the boundary value problem, if the boundary condition is of the first kind, $g(x, \xi)$ can be obtained by adding to a fundamental solution $-1/4\pi r$ the solution $\omega(x, \xi)$ of the following Dirichlet problem: $\Delta_x \omega(x, \xi) = 0$, $\omega(x, \xi) = 1/4\pi r$ ($x \in S$). We remark that there are slightly different definitions for Green's function. For example, there are cases where $g(x, \xi)$ is defined by $g(x, \xi) = 1/4\pi r + \omega(x, \xi)$ or by $g(x, \xi) = 1/r + \omega(x, \xi)$.

In the case in the previous paragraph, Green's functions satisfying the boundary conditions are uniquely determined. In general, if we are given a Green's function, then

for any regular function $v(x)$ the function

$$u(x) = \int_D g(x, \xi) v(\xi) d\xi$$

represents the solution of $L[u] = v$ with the boundary condition B . More precisely, if $v(x)$ satisfies the Hölder condition $|v(x) - v(x')| \leq L|x - x'|^\alpha$ ($0 < \alpha \leq 1$) (L, α positive constants), then $u(x)$ is of class C^2 . Conversely, if $u(x)$ satisfies the equation $L[u] = v$ and the boundary condition B , it is represented by the formula for $u(x)$. This means that if we denote the operator that associates u to v by G , then G is the inverse operator of the Laplacian L with the boundary condition B , and Green's function is the integral kernel of the operator G . Using this property, the boundary value problem relative to L can be reduced to a problem of integral equations. For example, the differential equation with the boundary condition B containing the complex parameter λ , $L[u] + \lambda u = f$, is equivalent to the integral equation $u + \lambda G[u] = G[f]$, which is obtained by letting G act from the left on the above differential equation. In this way, the problem can be simplified.

In the case of general boundary value problems for higher-order elliptic operators, Green's functions are defined in the same way as before (\rightarrow 189 Green's Operator). The important case is when L and the boundary condition B define a self-adjoint operator. In this case, Green's function is symmetric ($g(x, \xi) = g(\xi, x)$). To obtain Green's function is not easy in general. However, in some cases such functions can be obtained fairly easily (\rightarrow Appendix A, Table 15.VI).

B. Self-Adjoint Ordinary Differential Equations of the Second Order

Consider the operator $L[u] \equiv (p(x)u')' + q(x)u$ ($p(x) > 0$) defined in the interval $a \leq x \leq b$, with boundary conditions of the form $\alpha u' + \beta u = 0$ at the two endpoints. Then Green's function $g(x, \xi)$ is defined in the following way: (i) For $x \neq \xi$, $L[g(x, \xi)] = 0$; (ii) $[\partial g(x, \xi)/\partial x]_{x=\xi}^{\pm} = 0 = 1/p(\xi)$; (iii) for ξ fixed, $g(x, \xi)$ satisfies the homogeneous boundary conditions at $x = a$ and $x = b$.

Conditions (i) and (ii) mean that $L[g(x, \xi)] = \delta(x - \xi)$. We can construct $g(x, \xi)$ in the following way: Let $u_1(u_2)$ be the solution of $L[u] = 0$ satisfying the boundary condition at $x = a$ (at $x = b$). If u_1 and u_2 are linearly independent, we can satisfy $p(u_1' u_2 - u_1 u_2') = 1$ by choosing the constants suitably. Then Green's function $g(x, \xi)$ is defined by $g(x, \xi) = u_1(x)u_2(\xi)$ for $x \leq \xi$, and $g(x, \xi) = u_1(\xi)u_2(x)$ for $\xi \leq x$. If u_1 and u_2 are linearly dependent, there exists no

Green's function that satisfies conditions (i), (ii), and (iii). However, by modifying the definition, we can get a generalized Green's function playing a similar role [2]. This method can be applied to the case of ordinary differential equations of higher order.

C. The Laplace Operator

When the domain D is the n -dimensional sphere of radius a with center at the origin, Green's function of the Laplacian relative to the boundary condition $u = 0$ is obtained in the following way. Let $E(r)$ be the following fundamental solution of the Laplacian: $E(r) = (2\pi)^{-1} \log r$ for $n = 2$ and $E(r) = -((n-2) \cdot \omega_n r^{n-2})^{-1}$ for $n \geq 3$, where $\omega_n = 2\pi^{n/2} \Gamma(n/2)$ is the $(n-1)$ -dimensional surface area of the n -dimensional unit sphere. Then Green's function $g(x, \xi)$ is defined by

$$g(x, \xi) = E(r) - E(\rho r'/a),$$

where $\rho = (\sum_{i=1}^n \xi_i^2)^{1/2}$, $r' = (\sum_{i=1}^n (x_i - \xi_i)^2)^{1/2}$, $\xi'_i = (a/\rho)^2 \xi_i$.

D. Helmholtz's Differential Equation

Let D be an exterior domain with a smooth boundary S in \mathbf{R}^3 . In mathematical physics, the boundary value problem of finding a solution $u(x)$ of Helmholtz's differential equation $(\Delta + k^2)u(x) = f(x)$ ($k > 0$) satisfying $u(x) = 0$ ($x \in S$) is of particular interest. In this case, concerning the behavior of $u(x)$ at infinity, we usually assume Sommerfeld's radiation condition:

When $|x| \rightarrow +\infty$, $u(x) = O(|x|^{-1})$,

$$\left(\frac{\partial}{\partial r} u - iku \right) \Big|_{|x|=r} = o(|x|^{-1}),$$

where $\partial/\partial r$ is the derivative along the radial direction. It is known that this condition ensures the uniqueness of the solution (Rellich's uniqueness theorem). We can construct Green's function $G(x, \xi)$ for any $k(>0)$ so that for smooth $f(x)$ with bounded support,

$$u(x) = \int_D G(x, \xi) f(\xi) d\xi$$

represents the solution satisfying $u(x) = 0$ ($x \in S$) and Sommerfeld's radiation condition. Then, with

$$G(x, \xi) = -\frac{e^{ik|x-\xi|}}{4\pi|x-\xi|} + K_c(x, \xi),$$

where

$$|x - \xi| = \left(\sum_{i=1}^3 (x_i - \xi_i)^2 \right)^{1/2},$$

$K_c(x, \xi)$ can be obtained by solving an integral equation of \dagger Fredholm type [3, 4]. In this case, there exists no Green's operator in L_2 space, and $G(x, \xi)$ can be considered to be a generalized Green's function [4].

E. Stokes's Differential Equation

Let D be a bounded domain in \mathbf{R}^3 with smooth boundary S , and consider **Stokes's differential equation** in D

$$\mu \Delta u_i = \frac{\partial p}{\partial x_i} - \rho X_i, \quad i = 1, 2, 3, \quad \sum_{j=1}^3 \frac{\partial u_j}{\partial x_j} = 0,$$

where μ, ρ are positive constants. In hydrodynamics, we consider the boundary value problem of finding solutions $(u_1(x), u_2(x), u_3(x), p(x))$ of Stokes's equation satisfying the boundary condition $u_i(x) = 0, x \in S (i = 1, 2, 3)$. In this case, **Green's tensors** $G_{ij}(x, \xi), g_j(x, \xi)$ can be constructed, and for smooth functions $X_i(x) (i = 1, 2, 3)$ the unique solution of this boundary value problem is represented by

$$u_i(x) = \rho \sum_{j=1}^3 \int_D G_{ij}(x, \xi) X_j(\xi) d\xi,$$

$$p(x) = \rho \sum_{j=1}^3 \int_D g_j(x, \xi) X_j(\xi) d\xi$$

[5].

F. Parabolic Equations

Consider the boundary value problem (the initial boundary value problem):

$$L[u] \equiv \frac{\partial u}{\partial t} - c^2 \frac{\partial^2 u}{\partial x^2} = f(x, t), \quad t > 0, \quad a < x < b,$$

$$u(x, 0) = \varphi(x),$$

where at $x = a$ and $x = b$, $u(x, t)$ satisfies some homogeneous boundary conditions.

In this case, we can construct the function $g(x, t; \xi, \tau) (t \geq \tau)$ satisfying the following conditions: (i) $L[g] = 0$ except for $x = \xi, t = \tau$;

$$g(x, t; \xi, \tau) = \frac{\exp(-(x-\xi)^2/4c^2(t-\tau))}{2c\sqrt{\pi(t-\tau)}} + (\text{regular function})$$

in a neighborhood of $x = \xi, t = \tau$; and (iii) $g(x, t; \xi, \tau)$ satisfies the given homogeneous boundary conditions at $x = a$ and $x = b$. Then

$$u(x, t) = \int_0^t \int_a^b g(x, t; \xi, \tau) f(\xi, \tau) d\xi d\tau + \int_a^b g(x, t; \xi, 0) \varphi(\xi) d\xi$$

represents the solution of the problem stated

in this section for regular functions f, φ . The function $g(x, t; \xi, \tau)$ is called **Green's function** relative to the boundary value problem. Detailed consideration of such elementary cases is found in [6].

G. Kernel Functions

The kernel function is closely related to Green's function of Δ (the Laplacian) relative to the first boundary value problem in a domain in \mathbf{R}^2 .

First we explain the general definitions of the kernel function. Let E be a general set, and let \mathfrak{F} be a \dagger Hilbert space of complex-valued functions defined on E with a suitable inner product (f, g) . Suppose that we are given a function $K(x, y)$ defined on $E \times E$ satisfying the following conditions: (i) For any fixed y , $K(x, y)$ regarded as a function of x belongs to \mathfrak{F} ; and (ii) for any $f(x) \in \mathfrak{F}, (f(x), K(x, y))_x = f(y)$. Then $K(x, y)$ is called a **kernel function** or **reproducing kernel**. The kernel function, if it exists, is unique and is \dagger positive definite Hermitian; that is,

$$\sum_{j,k=1}^n K(y_j, y_k) \xi_j \bar{\xi}_k \geq 0. \quad (1)$$

Conversely, any positive definite function is a reproducing kernel of some Hilbert space. A necessary and sufficient condition for the existence of the kernel function is that for any $y \in E$, the linear functional $f \rightarrow f(y)$ be bounded. In this case, the minimum of $\|f\|$ under the condition $f(y) = 1 (f \in \mathfrak{F})$ is attained by the element $K(x, y)/K(y, y)$, and its value is $K(y, y)^{-1/2}$. When \mathfrak{F} is a \dagger separable space, then by an \dagger orthonormal system $\{\varphi_v(x)\}$, we can represent $K(x, y)$ as

$$K(x, y) = \sum_{v=1}^{\infty} \varphi_v(x) \overline{\varphi_v(y)}. \quad (2)$$

As an example of kernel functions, the following case is of particular importance. Let E be an n -dimensional \dagger complex manifold, φ and ψ be holomorphic \dagger differential forms of degree n on E , and let the following inner product be given:

$$(\varphi, \psi) = \int_E \varphi \wedge \bar{\psi}.$$

Now \mathfrak{F} is given by $\mathfrak{F} = \{\varphi \mid \varphi, \varphi < +\infty\}$, and the kernel function is called the **kernel differential**. When E is a domain in \mathbf{C}^n , regarding the coefficients of the differential form as functions, we call the kernel function **Bergman's kernel function**. Moreover, if E can be mapped onto a bounded domain by a one-to-one holomorphic mapping, then

$$ds^2 = \sum (\partial^2 \log K(z, \bar{z}) / \partial z_j \partial \bar{z}_k) dz_j d\bar{z}_k$$

is positive definite and gives a Kähler metric which is called the **Bergman metric**.

H. Kernel Functions for Domains in the Complex Plane

Let E be a domain D in the complex plane ($z = x + iy$). Let $K(z, \zeta)$ be Bergman's kernel function of D , and let $G(z, \zeta)$ be Green's function of Δ relative to the first boundary condition with a pole at ζ . Then we have

$$K(z, \zeta) = -(2/\pi) \partial^2 G(z, \zeta) / \partial z \partial \bar{\zeta}. \quad (3)$$

Next, let $U(z, \zeta)$ be the kernel function of the Hilbert space consisting of the holomorphic differential forms whose integrals are single-valued, and let $N(z, \zeta)$ be **Neumann's function** of Δ , i.e., the function that is harmonic in $D - \{\zeta\}$, has the same singularity as G at ζ , and whose derivative in the normal direction $\partial N / \partial n$ is constant along the boundary. Then we have

$$U(z, \zeta) = (2/\pi) \partial^2 N(z, \zeta) / \partial z \partial \bar{\zeta}. \quad (4)$$

Now the kernel $H(z, \zeta) = (N(z, \zeta) - G(z, \zeta)) / 2\pi$ is the kernel function relative to the Hilbert space consisting of all real harmonic functions whose integral mean value along the boundary Γ is 0 and having the inner product

$$(\varphi, \psi) = \iint_D \left(\frac{\partial \varphi}{\partial x} \frac{\partial \psi}{\partial x} + \frac{\partial \varphi}{\partial y} \frac{\partial \psi}{\partial y} \right) dx dy.$$

The kernel $H(z, \zeta)$ is called a **harmonic kernel function**.

Suppose that the boundary Γ is piecewise smooth, and consider the space of all holomorphic functions in D that are continuous on the boundary of D . The inner product of such functions φ and ψ is given by $(\varphi, \psi) = \int_{\Gamma} \varphi \bar{\psi} ds$ (ds is the element of the arc length of Γ). Hence we have a Hilbert space. Then the kernel function relative to this Hilbert space is called **Szegő's kernel function**, which has a close relation with bounded functions.

The kernel functions enable us to represent holomorphic mappings that map the domain D onto various canonical domains (\rightarrow 77 Conformal Mappings).

References

- [1] R. Courant and D. Hilbert, *Methods of mathematical physics*, Interscience, I, 1953; II, 1962.
- [2] K. Yosida, *Lectures on differential and integral equations*, Interscience, 1960. (Original in Japanese, 1950.)
- [3] V. D. Kupradze, *Randwertaufgaben der Schwingungstheorie und Integralgleichungen*,

Deutscher Verlag der Wiss., 1956. (Original in Russian, 1950.)

- [4] S. Mizohata, *Theory of partial differential equations*, Cambridge Univ. Press, 1973. (Original in Japanese, 1965.)
 - [5] F. K. G. Odqvist, Über die Randwertaufgaben der Hydrodynamik zäher Flüssigkeiten, *Math. Z.*, 32 (1930), 329–375.
 - [6] E. E. Levi, Sull' equazione del calore, *Ann. Mat. Pura Appl.*, (3) 14 (1908), 187–264.
- For kernel functions,
- [7] S. Bergman, The kernel function and conformal mapping, *Amer. Math. Soc. Math. Surveys*, 1950.
 - [8] S. Bergman and M. Schiffer, *Kernel functions and elliptic differential equations in mathematical physics*, Academic Press, 1953.
 - [9] N. Aronszajn, Theory of reproducing kernels, *Trans. Amer. Math. Soc.*, 68 (1950), 337–404.
 - [10] H. Meschkowski, *Hilbertsche Räume mit Kernfunktion*, Springer, 1962.

189 (XIII.29) Green's Operator

A. General Remarks

Consider the first and the third boundary value problems for the elliptic equation

$$A[u] \equiv -\Delta u + \sum_{i=1}^n a_i(x) \frac{\partial u}{\partial x_i} + c(x)u = f(x)$$

(\rightarrow 323 Partial Differential Equations of Elliptic Type). Let D be a bounded domain of \mathbf{R}^n whose boundary S consists of a finite number of smooth hypersurfaces. By the method of orthogonal projection, we take the domain $\mathscr{D}(A)$ as follows: (i) $\{u(x) | u(x) \in H^2(D) \text{ and } u(x) = 0 \text{ for } x \in S\}$ or (ii) $\{u(x) | u(x) \in H^2(D) \text{ and } \partial u / \partial n + \beta(x)u = 0 \text{ for } x \in S\}$ according as we are considering the first or third boundary value problem, where $H^2(D)$ is the Sobolev space (\rightarrow 168 Function Spaces). If the operator A is a one-to-one mapping from $\mathscr{D}(A)$ onto the function space $L_2(D)$, we call the inverse operator A^{-1} **Green's operator** relative to the boundary condition, and we denote it by G . In general, the existence of A^{-1} is not guaranteed. However, if we take real t large enough, $G_t = (A + tI)^{-1}$ exists.

Consider the general case where λ is a complex parameter: $(\lambda I - A)[u] = f(x)$, $f(x) \in L_2(D)$. Letting G_t act from the left, we have $(I - (\lambda + t)G_t)[u] = -G_t f$. Conversely, if $u \in L_2(D)$ is a solution, clearly $u(x) \in \mathscr{D}(A)$, and $u(x)$ satisfies the first partial differential equation and the boundary condition. Since G_t is a com-

compact operator in $L_2(D)$, the \dagger Riesz-Schauder theorem can be applied (\rightarrow 68 Compact and Nuclear Operators). In particular, if $\lambda + t$ is not an \dagger eigenvalue of G_t , $u(x) = (\lambda I - A)^{-1} f = -(\lambda I - (\lambda + t)G_t)^{-1} G_t f$ represents a unique solution.

In the equations in the first paragraph of this section, if $a_i(x) = 0$ and $c(x)$ and $\beta(x)$ are real, then G_t is a \dagger self-adjoint operator in $L_2(D)$, and therefore the \dagger Hilbert-Schmidt expansion theorem can be applied. Namely, let $\{\lambda_i\}$ be the eigenvalues of A such that $A\omega_i(x) = \lambda_i\omega_i(x)$, where $\{\omega_i(x)\}$ is an \dagger orthonormal system in $L_2(D)$. Then for any $f(x) \in L_2(D)$, $f(x) = \sum_{i=1}^\infty (f, \omega_i)\omega_i(x)$, where the right-hand side is taken in the sense of \dagger mean convergence. Furthermore, for $f(x) \in \mathcal{D}(A)$, we have the expansion $(Af)(x) = \sum_{i=1}^\infty \lambda_i(f, \omega_i)\omega_i(x)$ in the same sense.

When G_t is not self-adjoint, let G_t^* be the \dagger adjoint operator of G_t in $L_2(D)$. Then G_t^* represents Green's operator relative to the equation

$$\begin{aligned} (A^* + tI)[v] &= -\Delta v - \sum_{i=1}^n \frac{\partial}{\partial x_i} (\overline{a_i(x)}v) \\ &\quad + (\overline{c(x)} + t)v \\ &= g(x), \end{aligned}$$

corresponding to the boundary conditions (i) $v(x) = 0$, $x \in S$ (first boundary condition) and (ii) $(\partial/\partial n)v + \beta'(x)v = 0$, $x \in S$, where

$$\beta'(x) = \overline{\beta(x)} + \sum_{i=1}^n \overline{a_i(x)} \cos nx_i,$$

with n the outer normal (third boundary condition) [2].

B. Elliptic Equations of Higher Order

Green's operator can be defined for elliptic equations of higher order. Consider the equation

$$\begin{aligned} A(x, \partial/\partial x)u(x) &= f(x), \quad x \in D; \\ B_j(x, \partial/\partial x)u(x) &= 0, \quad x \in S, \\ j &= 1, 2, \dots, b \ (b = m/2), \quad (1) \end{aligned}$$

where A is an \dagger elliptic operator of order m and the boundary operators $\{B_j\}$ satisfy: (i) At every point x of S , the normal direction is not \dagger characteristic for any B_j ; and (ii) the order m_j of B_j is less than m , and $m_j \neq m_k$ ($j \neq k$). The domain $\mathcal{D}(A)$ of A is defined by

$$\begin{aligned} \mathcal{D}(A) &= \{u(x) | u \in H^m(D) \text{ and} \\ &\quad B_j(x, \partial/\partial x)u(x) = 0 \text{ for } x \in S, \\ &\quad j = 1, 2, \dots, b\}. \end{aligned}$$

When A is a one-to-one mapping from $\mathcal{D}(A)$ onto $L_2(D)$, the inverse $G = A^{-1}$ is called **Green's operator**.

This general boundary value problem, especially the existence theorem, was treated under some algebraic conditions on A and $\{B_j\}$ by M. Schechter [5] who showed that $G[u] \in H^m(D)$ if $u(x) \in L_2(D)$ and that $G[u]$ depends continuously on u . In particular, if $m > n/2$, then by \dagger Sobolev's theorem, G is a continuous mapping from $L_2(D)$ into $C^0(\overline{D})$, and G is represented by an \dagger integral operator of Hilbert-Schmidt type (L. Gårding [6]). Namely, for any $f(x) \in L_2(D)$,

$$\begin{aligned} (Gf)(x) &= \int_D G(x, \xi) f(\xi) d\xi, \\ \iint |G(x, \xi)|^2 dx d\xi &< +\infty. \end{aligned}$$

In general, the function $G(x, \xi)$, obtained by the **kernel representation** of Green's operator G , is called **Green's function**.

On the other hand, consider, for example, the \dagger Dirichlet problem of Δ in \mathbb{R}^3 . Then Green's function is defined in the following way (\rightarrow 188 Green's Functions): $G(x, \xi) = -(4\pi|x - \xi|)^{-1} + u(x, \xi)$, where $u(x, \xi)$ satisfies (i) $\Delta_x u(x, \xi) = 0$, and (ii) $G(x, \xi)|_{x \in S} = 0$. The function defined in this manner coincides with Green's function defined as a kernel representation of Green's operator [1, 2].

Suppose that in problem (1) the \dagger elliptic operator $A(x, \partial/\partial x)$ is independent of x , and let $E(x)$ be a \dagger fundamental solution, i.e., $E(x)$ is a distribution solution of $A(\partial/\partial x)E(x) = \delta(x)$ ($\delta(x)$ is \dagger Dirac's δ -function; \rightarrow 112 Differential Operators). Then $E(x)$ is a C^∞ -function away from the origin. Moreover, in a neighborhood of the origin the following estimates hold: For $|\alpha| < m$,

$$\left| \left(\frac{\partial}{\partial x} \right)^\alpha E(x) \right| \leq \begin{cases} c|x|^{m-n-|\alpha|}, & m-n-|\alpha| < 0, \\ c \log|x|^{-1}, & m-n-|\alpha| = 0, \\ c, & m-n-|\alpha| > 0, \end{cases}$$

where c is some positive constant. When problem (1) has Green's operator G , we can say the following, using the fundamental solution: Green's function $G(x, \xi)$ exists and can be written as $G(x, \xi) = E(x - \xi) + u(x, \xi)$, where for any fixed $\xi \in D$, $u(x, \xi)$ satisfies (i) $A(\partial/\partial x)u(x, \xi) = 0$ and (ii) $B_j(x, \partial/\partial x)G(x, \xi) = 0$, $x \in S$, $j = 1, 2, \dots, b$.

C. Hypoelliptic Operators

Let

$$\begin{aligned} A(x, \partial/\partial x) &= \sum_{|\alpha| \leq m} a_\alpha(x) \left(\frac{\partial}{\partial x} \right)^\alpha, \\ \left(\frac{\partial}{\partial x} \right)^\alpha &= \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}, \end{aligned}$$

be a general partial differential operator with C^∞ -coefficients. If the kernel $E(x, \xi)$ satisfies $A(x, \partial/\partial x)E(x, \xi) = \delta(x - \xi)$, that is, if we have the relation $\langle E(x, \xi), {}^1A(x, \partial/\partial x)\varphi(x) \rangle_x = \varphi(\xi)$ for any $\varphi(x) \in \mathcal{D}$, then $E(x, \xi)$ is called a **fundamental solution** of A , where 1A is the **transposed operator** of A :

$${}^1A(x, \partial/\partial x)v(x) = \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \left(\frac{\partial}{\partial x} \right)^\alpha (a_\alpha(x)v(x)).$$

Now if there exists a fundamental solution $E'(x, \xi)$ of ${}^1A(x, \partial/\partial x)$ such that (i) $E'(x, \xi)$ defines a kernel that gives rise to two continuous mappings, one of which maps the space \mathcal{D}_ξ into \mathcal{E}_x , and the other of which maps the space \mathcal{D}_x into \mathcal{E}_ξ (\rightarrow 168 Function Spaces), and (ii) for $x \neq \xi$, $E'(x, \xi)$ a C^∞ -function of (x, ξ) , then any distribution $u(x)$ satisfying $A(x, \partial/\partial x)u(x) = g(x)$ is a C^∞ -function, where $g(x)$ is of class C^∞ . In general, an operator A with the property that any solution $u(x)$ of $A(x, \partial/\partial x)u(x) = g(x)$ is of class C^∞ whenever $g(x)$ is of class C^∞ , is called **hypoelliptic**. Elliptic and parabolic operators are both hypoelliptic. L. Hörmander characterized the hypoelliptic differential operators with constant coefficients [8] (\rightarrow 112 Differential Operators).

A kernel $E(x, \xi)$ such that

$$A(x, \partial/\partial x)E(x, \xi) = \delta(x - \xi) + \varpi(x, \xi),$$

with $\varpi(x, \xi)$ a C^∞ -function, is called a **parametrix** of A . To prove the hypoellipticity of A , it suffices to show the existence of a parametrix $E'(x, \xi)$ of the operator 1A having the properties (i) and (ii) mentioned in the previous paragraph.

To explain the notion of the fundamental solution for the evolution equations, suppose that we are given an evolution equation

$$L[u] = \frac{\partial^m}{\partial t^m} u(x, t) + \sum_{j < m} a_{x,j}(x, t) \left(\frac{\partial}{\partial x} \right)^j \frac{\partial^j}{\partial t^j} u(x, t) = 0,$$

$x \in \mathbf{R}^n$, $t_0 \leq t \leq T$. A kernel $E(x, t; \xi, t_0)$ ($t_0 \leq t \leq T$) is called a **fundamental solution** to the evolution equation if

$$L_{x,t}(E(x, t; \xi, t_0)) = 0, \quad t > t_0,$$

and

$$\lim_{t \rightarrow t_0 + 0} \frac{\partial^i}{\partial t^i} E(x, t; \xi, t_0) = \begin{cases} 0, & 0 \leq j \leq m-2, \\ \delta(x - \xi), & j = m-1. \end{cases}$$

References

- [1] H. G. Garnir, Les problèmes aux limites de la physique mathématique, Birkhäuser, 1958.
- [2] S. Mizohata, Theory of partial differential equations, Cambridge Univ. Press, 1973. (Original in Japanese, 1965.)

- [3] R. Courant and D. Hilbert, Methods of mathematical physics II, Interscience, 1962.
- [4] L. Gårding, Dirichlet's problem for linear elliptic differential equations, Math. Scand., 1 (1953), 55–72.
- [5] M. Schechter, General boundary value problems for elliptic partial differential equations, Comm. Pure Appl. Math., 12 (1959), 457–486.
- [6] L. Gårding, Applications of the theory of direct integrals of Hilbert spaces to some integral and differential operators, Univ. of Maryland lecture series no. 11, 1954.
- [7] L. Schwartz, Théorie des distributions, Hermann, second edition, 1966.
- [8] L. Hörmander, On the theory of general partial differential operators, Acta Math., 94 (1955), 161–248.

190 (IV.1) Groups

A. Definition

Let G be a nonempty set. Suppose that for any elements a, b of G there exists a uniquely determined element c of G , which is called the **product** of a and b , written $c = ab$. We call G a **group** or **multiplicative group** if (i) the **associative law** $a(bc) = (ab)c$ holds, and (ii) for any elements $a, b \in G$ there exist uniquely determined elements $x, y \in G$ satisfying $ax = b$, $ya = b$. Then the mapping $(a, b) \rightarrow ab$ is called **multiplication** in G . Condition (ii) is equivalent to the following two conditions: (iii) There exists an element e (called the **identity element** or **unit element** of G) such that $ae = ea = a$ for any element a of G ; and (iv) for any element a of G there exists an element x such that $ax = xa = e$.

The element x in condition (iv) is called the **inverse** (or **inverse element**) of a , denoted by a^{-1} . The uniqueness of the identity element e and the inverse a^{-1} follows readily from the axioms. The identity element of a multiplicative group is sometimes denoted by 1. If $ab = ba$, then we say that a and b **commute**. The **commutative law**, $ab = ba$ for any elements $a, b \in G$, is not assumed in general. A group satisfying the commutative law is called an **Abelian group** (or **commutative group**) in honor of N. H. Abel, who made use of commutative groups in his study of the theory of equations. The product in a commutative group is often written in the form $a + b$, and in this case the mapping $(a, b) \rightarrow a + b$ is called **addition**. The element $a + b$ is called the **sum** of a and b , and G is called an **additive group**. In an additive

group the identity element is usually denoted by 0 and called the **zero element**, and the inverse of a is denoted by $-a$ (\rightarrow 2 Abelian Groups; 277 Modules). To describe the †law of composition, we sometimes use notation different from multiplication or addition (\rightarrow 409 Structures).

B. Examples

A †linear space over a †field K is an additive group with respect to the usual addition of vectors (\rightarrow 256 Linear Spaces). A field is an additive group with respect to the addition, and the set of nonzero elements of a field forms a group with respect to the multiplication, which is called the **multiplicative group of the field** (\rightarrow 149 Fields).

All †invertible $n \times n$ matrices over a ring R form a group with respect to the usual multiplication of matrices. This group is called the †general linear group of degree n over R (\rightarrow 60 Classical Groups).

All one-to-one mappings from a set M onto itself (i.e., all **permutations** on M) form a group with respect to the composition defined by $(f \circ g)(x) = f(g(x))$ ($x \in M$). (Sometimes the product $f \circ g$ is denoted by gf and $f(x)$ by xf . Then $x(gf) = (xg)f$.) The group of all permutations on M is called the **symmetric group** on M . A group G is called a **permutation group** (on M) if every element of G is a permutation on M . For instance, the general linear group of degree n over a field K may be regarded as a permutation group on the set of n -dimensional vectors, and it is also regarded as a permutation group on a †tensor space.

All †motions in a Euclidean space form a group with respect to the usual composition of motions. All invertible $n \times n$ matrices over K leaving a given †quadratic form invariant form a group with respect to the usual multiplication of matrices. This group is called the †orthogonal group belonging to the given quadratic form. If K is the †complex number field or the real number field, then these groups are Lie groups (\rightarrow 13 Algebraic Groups; 151 Finite Groups; 161 Free Groups; 249 Lie Groups; 423 Topological Groups).

C. Fundamental Concepts

If a group G consists of a finite number of elements, then G is called a **finite group**; otherwise, G is called an **infinite group**. The number of elements of G is called the **order** of G . A nonempty subset H of G is called a **subgroup** of G if H is a group with respect to the multiplication of the group G . Hence a nonempty sub-

set H is a subgroup of G if and only if $a^{-1}b \in H$ for any $a, b \in H$. For a family $\{H_\lambda\}$ of subgroups of G , the intersection $\bigcap_\lambda H_\lambda$ is also a subgroup.

The associative law of multiplication says that elements a_1, a_2, a_3 of G determine the product $a_1 a_2 a_3$, which is the common value of $(a_1 a_2) a_3$ and $a_1 (a_2 a_3)$. This law can be generalized to say that any ordered set of n elements a_1, a_2, \dots, a_n ($n > 2$) of G determines their product $a_1 a_2 \dots a_n$ (**general associative law**).

When $a_1 = a_2 = \dots = a_n = a$, we denote the product $aa \dots a$ by a^n . If we define a^{-n} for $n \geq 0$ by $a^0 = e$ and $a^{-n} = (a^n)^{-1}$, we then have $a^n a^m = a^{n+m}$, $(a^n)^m = a^{nm}$ for any $n, m \in \mathbb{Z}$. If there exists a positive integer n such that $a^n = e$, then the smallest positive integer d with $a^d = e$ is called the **order** of the element a . If there is no such n , then a is called an element of **infinite order**. If a is of infinite order, then its powers $a^0 (= e), a^{\pm 1}, a^{\pm 2}, \dots$, are all unequal. If a is of order d , then the different powers of a are $a^0 (= e), a, a^2, \dots, a^{d-1}$. All the powers of a form a subgroup $\langle a \rangle$ of G , called a **cyclic subgroup**. The order of an element a is the same as the order of the subgroup $\langle a \rangle$. The group $\langle a \rangle$ itself is called a **cyclic group** and is an example of an Abelian group (\rightarrow 2 Abelian Groups).

Let S be a subset of a group G . Then the intersection of all subgroups of G containing S is called the subgroup **generated** by S and is denoted by $\langle S \rangle$. It is the smallest subgroup containing S , and if S is nonempty, $\langle S \rangle$ consists of all the elements of the form $a_1^{m_1} a_2^{m_2} \dots a_r^{m_r}$ ($a_i \in S, m_i \in \mathbb{Z}$). If $\langle S \rangle = G$, the elements of S are called **generators** of G . When G has a finite set of generators, G is said to be **finitely generated**. When $S = \{a\}$, then $\langle S \rangle$ coincides with $\langle a \rangle$, and the element a is the generator of the cyclic group $\langle a \rangle$. Suppose that elements a_1, \dots, a_n of G satisfy an equation of the form $a_1^{m_1} a_2^{m_2} \dots a_n^{m_n} = 1$. This equation is then called a **relation** among the elements a_1, \dots, a_n . If we have a system of generators and all relations among the generators, then they define a group (\rightarrow 161 Free Groups). It is, however, still an open problem to find a general procedure to decide whether the group determined by a given system of generators and the relations among them contains elements other than the identity (\rightarrow 161 Free Groups B).

For a subset S and an element x of a group G , the set of all elements $x^{-1}sx$ ($s \in S$) is denoted by $x^{-1}Sx$ or S^x , and S and S^x are called **conjugate**. We have $(ab)^x = a^x b^x$, $(a^{-1})^x = (a^x)^{-1}$. If H is a subgroup, then H^x is also a subgroup. For a subset S , the set of all elements x satisfying $S^x = S$ forms a subgroup $N(S)$, called the **normalizer** of S . The set of all elements that commute with every element of

S forms a subgroup $Z(S)$, called the **centralizer** of S . The centralizer Z of G is called the **center** of G . The set of all elements conjugate to a given element a of G is called a **conjugacy class**. A group G is the disjoint union of its conjugacy classes.

Let H be a subgroup of a group G and x an element of G . The set of elements of the form hx ($h \in H$) is denoted by Hx and is called a **right coset** of H . A **left coset** xH is defined similarly. G is the disjoint union of left (right) cosets of H . The cardinality of the set of left cosets of H equals that of the set of right cosets of H ; it is called the **index** of the subgroup H and is denoted by $(G:H)$. Given two subgroups H and K of G , the set $HxK = \{h x k \mid h \in H, k \in K\}$ is called the **double coset** of H and K , and G is the disjoint union of different double cosets of H and K . If the left cosets of a subgroup H are also the right cosets, i.e., if $Hx = xH$ for every $x \in G$, then H is called a **normal subgroup** (or **invariant subgroup**) of G . An equivalent condition is that $H = H^x$ for all $x \in G$. The center of G is always a normal subgroup of G . If H is a normal subgroup of G , the set of all products of an element of Ha and an element of Hb coincides with Hab . Thus if we define the product of two cosets Ha and Hb to be Hab , then the set of cosets of H forms a group. This group is denoted by G/H and is called the **factor group** (or **quotient group**) of G modulo H . (When G is an additive group, G/H is also denoted by $G - H$ and is called the **difference group**.) The group G itself and $\{e\}$ are normal subgroups of G . If G has no normal subgroup other than these two, then G is called a **simple group**. A subgroup of finite index contains a normal subgroup of finite index. If H is a subgroup of finite index, then we can find a common complete system of representatives of the left cosets and the right cosets of H . If G is finitely generated, then so is any subgroup of G of finite index.

Let R be an \dagger equivalence relation defined in a group G . If xRx' and yRy' always imply $(xy)R(x'y')$, then we say that R is **compatible** with the multiplication. The \dagger quotient set G/R is a group with respect to the induced multiplication. This group is called the **quotient group** of G with respect to R . The equivalence class H containing e is a normal subgroup, and xRx' if and only if $x^{-1}x' \in H$, i.e., x and x' are contained in the same coset of H . Thus G/R coincides with G/H .

If G is a finite group of order n , then the order and the index of any subgroup of G , the order of any element of G , the cardinal number of any conjugacy class of G , and the number of different conjugate subgroups of any subgroup of G are all divisors of n .

D. Isomorphisms and Homomorphisms

If there is a one-to-one mapping $a \leftrightarrow a'$ of the elements of a group G onto those of a group G' and if $a \leftrightarrow a'$ and $b \leftrightarrow b'$ imply $ab \leftrightarrow ab'$, then we say that G and G' are **isomorphic** and write $G \cong G'$. If we put $a' = f(a)$, then $f: G \rightarrow G'$ is a \dagger bijection satisfying $f(ab) = f(a)f(b)$ ($a, b \in G$). More generally, if a mapping $f: G \rightarrow G'$ satisfies $f(ab) = f(a)f(b)$ for all $a, b \in G$, then f is called a **homomorphism** of G to G' . An \dagger injective (\dagger surjective) homomorphism is also called a \dagger monomorphism (\dagger epimorphism). If there is a surjective homomorphism $G \rightarrow G'$, then we say that G' is **homomorphic** to G . The composite of two homomorphisms is also a homomorphism. If a homomorphism $f: G \rightarrow G'$ is a bijection, then f is called an **isomorphism**. In this case f^{-1} is also an isomorphism, and we have $G \cong G'$.

For a subgroup H of a group G , the injective homomorphism $f: H \rightarrow G$ defined by $f(a) = a$ ($a \in H$) is called the **canonical injection** (or **natural injection**). For a factor group G/R of G , the surjective homomorphism $f: G \rightarrow G/R$ such that $a \in f(a)$ ($a \in G$) is called the **canonical surjection** (or **natural surjection**).

Let $f: G \rightarrow G'$ be a homomorphism. Then the **image** $f(G)$ of f is a subgroup of G' , and the **kernel** $H = \{a \in G \mid f(a) = e' \text{ (the identity of } G')\}$ of f is a normal subgroup of G . The equivalence classes of the equivalence relation given by $f(x) = f(y)$ are just the cosets of H , and f induces an isomorphism $\bar{f}: G/H \rightarrow f(G)$. The latter proposition is called the **homomorphism theorem** of groups. This theorem is extended in the following way: For simplicity let $f: G \rightarrow G'$ be a surjective homomorphism. (i) If H' is a normal subgroup of G' , then the inverse image $H = f^{-1}(H')$ is a normal subgroup of G , and f induces the isomorphism $\bar{f}: G/H \rightarrow G'/H'$. (ii) if H is a subgroup and N is a normal subgroup of G , then $HN = \{hn \mid h \in H, n \in N\}$ is a subgroup of G , and the canonical injection $H \rightarrow HN$ induces an isomorphism $H/H \cap N \rightarrow HN/N$. (iii) If H and N are two normal subgroups of G such that $H \supset N$, then the canonical surjection $G \rightarrow G/N$ induces an isomorphism $G/H \rightarrow (G/N)/(H/N)$. Propositions (i), (ii), and (iii) are called the **isomorphism theorems** of groups.

A homomorphism of G to itself is called an **endomorphism** of G , and an isomorphism of G to itself is called an **automorphism** of G . The set of automorphisms of G forms a group with respect to the composition of mappings, called the **group of automorphisms** of G . Given an element a of G , the mapping $x \rightarrow a^{-1}xa$ ($x \in G$) yields an automorphism of G which is called an **inner automorphism** of G . The set of inner automorphisms of G forms a normal subgroup of the group of automorphisms of G , called the

group of inner automorphisms of G , which is isomorphic to the factor group of G modulo its center. The factor group of the group of automorphisms of G modulo the group of inner automorphisms of G is called the **group of outer automorphisms** of G .

If a mapping $f: G \rightarrow G'$ from a group G to another group G' satisfies $f(ab) = f(b)f(a)$ ($a, b \in G$), then f is called an **antihomomorphism**. A bijective antihomomorphism is called an **anti-isomorphism**. When $G = G'$, f is called an **anti-endomorphism** or **anti-automorphism** (e.g., $f: G \rightarrow G$ defined by $f(a) = a^{-1}$ is an anti-isomorphism).

E. Groups with Operator Domain

Let Ω be a set and G a group. Suppose that for each $\theta \in \Omega$ and $x \in G$, the product $\theta x \in G$ is defined and satisfies $\theta(xy) = \theta(x)(\theta y)$. Then Ω is called an **operator domain** of G , and G is called a group with operator domain Ω , or simply an **Ω -group**. (We sometimes write x^θ instead of θx .) The mapping $(\theta, x) \rightarrow \theta x$ from $\Omega \times G$ to G is called the \circ operation of Ω on G . If G is an Ω -group, then any element θ of Ω induces an endomorphism $\theta_G: x \rightarrow \theta x$ of G . Conversely, if we are given a mapping $\theta \rightarrow \theta_G$ of Ω to the set of endomorphisms of G , then we may regard G as an Ω -group. Any group may be regarded as an Ω -group with Ω equal to the empty set or to the set consisting of the identity automorphism of G . Thus the general theory of groups can be extended to the theory of groups with operator domain, and in some cases effective use of suitable operator domains can be fruitful in the investigation of the properties of groups themselves. (\rightarrow 2 Abelian Groups; 277 Modules).

A subgroup H of an Ω -group G is called an **Ω -subgroup** (or **admissible subgroup**) if $\theta x \in H$ for any $\theta \in \Omega$ and $x \in H$. In this case, H is also an Ω -group. If an equivalence relation R defined in G is compatible with the multiplication and also compatible with the operators, namely, if xRx' implies $(\theta x)R(\theta x')$ for any $\theta \in \Omega$, then the quotient group G/R is also an Ω -group. The equivalence class containing e is an **admissible normal subgroup**. Conversely, if H is an admissible normal subgroup, then the equivalence relation defined by H is compatible with the operators, and the factor group G/H is an Ω -group. A homomorphism $f: G \rightarrow G'$ of an Ω -group G to an Ω -group G' is called an **Ω -homomorphism** (**admissible homomorphism** or **operator homomorphism**) if $f(\theta x) = \theta f(x)$ for any $\theta \in \Omega$ and $x \in G$. If f is an isomorphism, then f is called an **Ω -isomorphism** (**admissible isomorphism** or **operator isomor-**

phism). We have the homomorphism theorem and the isomorphism theorems of Ω -groups if we consider only admissible subgroups and admissible homomorphisms.

F. Sequences of Subgroups

Let H_1, H_2, \dots be an infinite sequence of (normal) subgroups of a group G . If $H_i \subseteq H_{i+1}$ ($i = 1, 2, \dots$), then the sequence is called an **ascending chain** of (normal) subgroups. If $H_i \not\subseteq H_{i+1}$ ($i = 1, 2, \dots$), then it is called a **descending chain** of (normal) subgroups. If there is no ascending (or descending) chain of (normal) subgroups of G , we say that G satisfies the **ascending (or descending) chain condition** for (normal) subgroups. These conditions are the same as the ascending (or descending) chain condition in the ordered set of all (normal) subgroups of G (\rightarrow 311 Ordering C). A group G satisfies the ascending chain condition for subgroups if and only if every subgroup of G is finitely generated. Also, for groups with operator domain we have similar results. The structure of Abelian groups satisfying the ascending (descending) chain condition is completely determined (\rightarrow 2 Abelian Groups). It is not known whether there is an infinite group satisfying both the ascending and descending chain conditions for subgroups. A group satisfying the descending chain condition for subgroups has no element of infinite order, but the converse is not true. There is an infinite group which is finitely generated and has no element of infinite order (\rightarrow 161 Free Groups C).

G. Normal Chains

A finite sequence $G = G_0 \supset G_1 \supset G_2 \supset \dots \supset G_r = \{e\}$ of subgroups of a group G is called a **normal chain** if G_i is a normal subgroup of G_{i-1} for $i = 1, 2, \dots, r$. We call r the **length** of the chain. The sequence $G_0/G_1, G_1/G_2, \dots, G_{r-1}/G_r$ is called the **sequence of factor groups** of the normal chain. A normal chain $G = H_0 \supset H_1 \supset H_2 \supset \dots \supset H_s = \{e\}$ is called a **refinement** of the chain $G = G_0 \supset G_1 \supset \dots \supset G_r = \{e\}$ if every G_i appears in this chain. Two normal chains with the same length are called isomorphic if there is a one-to-one correspondence between their sequences of factor groups such that corresponding factor groups are isomorphic. Any two normal chains have refinements which are isomorphic to each other (Schreier's refinement theorem). A normal chain is called a **composition series** (or **Jordan-Hölder sequence**) if it consists of different subgroups of G and in

any proper refinement there appear two successive subgroups which are the same. The sequence of factor groups of a composition series is called a **composition factor series**, and the factor groups appearing in this series are called **composition factors**. Any composition factor is a simple group. As a direct consequence of the refinement theorem we see that if a group G has a composition series, then the composition factor series is unique up to isomorphism and the ordering of the factors. (This theorem is due to O. Hölder. C. Jordan proved that if G is a finite group, then the set of orders of composition factors is independent of the choice of composition series. Hence we call the theorem the **Jordan-Hölder theorem**.)

For an Ω -group G , if we consider only Ω -subgroups, we have definitions and theorems similar to those in this section. When we take the group of inner automorphisms of G as Ω , then a composition series of the Ω -group G is called a **principal series**. If we take the group of automorphisms of G as Ω , then a composition series is called a **characteristic series**. An infinite group G does not always have a composition series. Even if G has a composition series, G may have an infinite normal chain $G_1 \subset G_2 \subset \dots \subset G$ such that each G_i is a normal subgroup of G_{i+1} and $\bigcup G_i = G$. In fact, there is a simple group which has such an infinite normal chain (P. Hall). Two groups which have isomorphic composition series are not necessarily isomorphic. A subgroup of a group G is called a **subnormal subgroup** of G if it may appear in some normal chain. The intersection of two subnormal subgroups is also subnormal, but their **join** (i.e., the subgroup generated by both of them) is not necessarily subnormal in an infinite group. The set of subgroups and the set of normal subgroups of a group form \dagger lattices with respect to the inclusion relation (for the relationship between these lattices and the group structure \rightarrow [8]).

H. Commutator Subgroups

Given two elements a and b of a group G , we call $a^{-1}b^{-1}ab = [a, b]$ the **commutator** of a and b . The subgroup C generated by all commutators in G is called the **commutator subgroup** (or **derived group**) of G . The subgroup C is a normal subgroup of G , and the factor group G/C is Abelian. On the other hand, if B is a normal subgroup of G and G/B is Abelian, then B contains C . For two subsets A, B of G , the subgroup generated by the commutators $[a, b]$ ($a \in A, b \in B$) is called the **commutator group of A and B** and is denoted by $[A, B]$. If A and B are normal subgroups of G , then $C' = [A, B]$ is

also a normal subgroup of G , and A/C' commutes with B/C' elementwise in the factor group G/C' . Furthermore, $[A, B]$ is the minimal normal subgroup with the property. The subgroup $[G, G]$ is the commutator subgroup of G .

If the commutator subgroup of G is Abelian, then G is called a **meta-Abelian group**. If a group G has a normal chain $G (= G_0) \supset G_1 \supset G_2 (= \{e\})$ of length 2 and the factor groups $G/G_1, G_1/G_2$ are Abelian, then G is meta-Abelian. Meta-Abelian groups are special cases of solvable groups, discussed in Section I.

I. Solvable Groups

Suppose that we are given a series of subgroups G_i ($i = 0, 1, 2, \dots$) of G such that $G = G_0$ and $[G_i, G_i] = G_{i+1}$. Then we have a normal chain $G = G_0 \supset G_1 \supset G_2 \supset \dots$. If $G_r = \{e\}$ for some r , then G is called a **solvable group**. For the normal chain $G (= G_0) \supset G_1 \supset \dots \supset G_r (= \{e\})$ the factor groups G_i/G_{i+1} ($i = 0, 1, \dots, r-1$) are all Abelian. A finite group G is solvable if and only if G has a composition series $G = H_0 \supset H_1 \supset H_2 \supset \dots \supset H_s = \{e\}$ such that the factor groups H_i/H_{i+1} ($i = 0, 1, \dots, s-1$) are all of prime order. An \dagger irreducible algebraic equation over a field of \dagger characteristic 0 is solvable by radicals if and only if its Galois group is solvable (\rightarrow 172 Galois Theory).

J. Nilpotent Groups

The sequence of subgroups $G = G_0 \supset G_1 \supset G_2 \supset \dots$ defined inductively by setting $G_r = [G, G_{r-1}]$ ($r = 1, 2, \dots$) is called the **lower central series** of G . If $G_n = \{e\}$ for some n , then G is called a **nilpotent group**, and the least number n with $G_n = \{e\}$ is called the **class** of the nilpotent group G . A nilpotent group is solvable. Let Z_1 be the center of G , Z_2/Z_1 be the center of G/Z_1 , and so on. Then we have a sequence of subgroups $Z_0 = \{e\} \subset Z_1 \subset Z_2 \subset \dots$, called the **upper central series** of G . A group G is nilpotent if and only if $Z_m = G$ for some m , and the least number m with $Z_m = G$ is the class of G . For the subgroups G_r and Z_r ($r = 1, 2, \dots$), we have $[G_{i-1}, Z_i] = \{e\}$. If G is a \dagger Lie group, then G is nilpotent if and only if the corresponding \dagger Lie algebra \mathfrak{g} is nilpotent, i.e., $\mathfrak{g}^n = 0$.

K. Infinite Solvable Groups

The concepts of solvability and nilpotency are generalized in several ways for infinite groups. For instance, a group G is called a **generalized solvable group** if any homomorphic image of G

which is unequal to $\{e\}$ contains an Abelian normal subgroup unequal to $\{e\}$, and G is called a **generalized nilpotent group** if any homomorphic image ($\neq \{e\}$) of G has center unequal to $\{e\}$. These definitions coincide with the previous ones for finite groups but not for infinite groups [7].

L. Direct Products

Let G_1, \dots, G_n be a finite number of groups. The set G of all elements (x_1, \dots, x_n) with $x_i \in G_i$ ($i = 1, \dots, n$) is a group if we define the product of two elements $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ to be $xy = (x_1 y_1, \dots, x_n y_n)$. We call G the **direct product** of groups G_1, \dots, G_n and write $G = G_1 \times \dots \times G_n$. If e_j is the identity element of G_j , then $e = (e_1, \dots, e_n)$ is the identity element of G . The mapping $(x_1, \dots, x_n) \rightarrow x_i$ from G to G_i is a surjective homomorphism, called the **canonical surjection**. The subgroup $H_i = \{(e_1, \dots, e_{i-1}, x_i, e_{i+1}, \dots, e_n) \mid x_i \in G_i\}$ is isomorphic to G_i . The subgroups H_i ($i = 1, \dots, n$) satisfy the following conditions: (i) H_i is a normal subgroup of G . (ii) H_i commutes with H_j elementwise if $i \neq j$. (iii) Any element of G can be written uniquely as the product of elements of H_1, \dots, H_n . Conversely, if a group G has subgroups H_1, \dots, H_n satisfying these three conditions, then G is isomorphic to $H_1 \times \dots \times H_n$. In this case we also write $G = H_1 \times \dots \times H_n$, and we call this a **direct decomposition** of G . Each H_i is called a **direct factor** of G . Conditions (i), (ii), and (iii) are equivalent to condition (i), (ii') $G = H_1 H_2 \dots H_n$, and $H_1 \dots H_{i-1} \cap H_i = \{e\}$ ($i = 2, \dots, n$).

A group G is called **indecomposable** if G cannot be decomposed into the direct product of two subgroups unequal to $\{e\}$, and **completely reducible** if G is the direct product of simple groups. If G satisfies the ascending or descending chain condition for normal subgroups, then G can be decomposed into the direct product of indecomposable groups. Such a decomposition is not unique in general, but if G has two direct product decompositions $G = G_1 \times \dots \times G_m = H_1 \times \dots \times H_n$, where G_i and H_j are indecomposable and not equal to $\{e\}$, then $m = n$ and the factors G_i are isomorphic to the factors H_j for some j ; moreover, if G_1 corresponds, say, with H_1 , then we have $G = H_1 \times G_2 \times \dots \times G_m$. This fact was first stated by J. H. M. Wedderburn, and a complete proof of the theorem was given by R. Remak and O. Schmidt. Later W. Krull extended it to more general groups (with operator domain), and we call it the **Krull-Remak-Schmidt theorem**. O. Ore formulated it as a theorem on π -modular lattices.

For an infinite number of groups G_λ ($\lambda \in \Lambda$)

we define the direct product $\prod_{\lambda \in \Lambda} G_\lambda$ of these groups similarly. The set of all elements $(\dots, x_\lambda, \dots)$ ($x_\lambda \in G_\lambda$) such that almost all x_λ (i.e., all except a finite number of λ) are identity elements is a subgroup of the direct product, called the **direct sum** (or **restricted direct product**) of $\{G_\lambda\}$.

M. Free Products

Given a family of groups $\{G_\lambda\}_{\lambda \in \Lambda}$, we define the most general group G generated by these groups, called the **free product** of $\{G_\lambda\}$, together with canonical injections $f_\lambda: G_\lambda \rightarrow G$.

Let S be the disjoint union of the sets $\{G_\lambda\}_{\lambda \in \Lambda}$, and regard G_λ as a subset of S . A **word** is either void or a finite sequence a_1, a_2, \dots, a_n of elements of S , and we denote the set of all words by W . The product of two words w and w' is defined by connecting w with w' so that the π -associative law holds. We write $w \succsim w'$ when two words w and w' satisfy one of the following two conditions: (i) The word w has successive members a, b which belong to the same group G_λ , and the word w' is obtained from w by replacing a, b by the product ab . (ii) Some member of w is an identity element, and w' is the word obtained from w by deleting this member. For two words w and w' , we write $w \equiv w'$ if there is a finite sequence of words $w = w_0, w_1, \dots, w_n = w'$ such that for each i ($1 \leq i \leq n$), either $w_{i-1} \succsim w_i$ or $w_i \succsim w_{i-1}$. This relation is an equivalence relation and is compatible with the multiplication. Thus we may define a multiplication for the quotient set G of W by this equivalence relation, and then G is a group whose identity element is the equivalence class containing the void word. Any $x \in G_\lambda$ is regarded as a word, and we have an injective homomorphism $f_\lambda: G_\lambda \rightarrow G$ by assigning the corresponding class to each element of G_λ . The group G is called the **free product** of the system of groups $\{G_\lambda\}_{\lambda \in \Lambda}$, and f_λ is called the **canonical injection**. The free product G of $\{G_\lambda\}_{\lambda \in \Lambda}$ is characterized by the following universal property: Given a group G' and homomorphisms $f'_\lambda: G_\lambda \rightarrow G'$ ($\lambda \in \Lambda$), we can find a unique homomorphism $g: G \rightarrow G'$ such that $g \circ f_\lambda = f'_\lambda$. The free product is the dual concept of direct product and is also called the π -coproduct (\rightarrow 52 Categories and Functors). If each G_λ is an infinite cyclic group generated by a_λ , then the free product of the G_λ is the π -free group generated by $\{a_\lambda\}$ (\rightarrow 161 Free Groups).

The concept of free product is generalized in the following way. Let H be a fixed group. We consider the family of pairs (G, j) , where G is a group and $j: H \rightarrow G$ is an injective homomorphism. A homomorphism of pairs $f: (G, j) \rightarrow$

(G', j') is defined to be a group homomorphism $f: G \rightarrow G'$ such that $f \circ j = j'$. For a given family of pairs $\{(G_\lambda, j_\lambda)\}$, we have the **amalgamated product** (G, j) of the family and the canonical homomorphism $f_\lambda: (G_\lambda, j_\lambda) \rightarrow (G, j)$, which is characterized by the following universal property: Given a pair (G', j') and homomorphisms $f'_\lambda: (G_\lambda, j_\lambda) \rightarrow (G', j')$, we have a unique homomorphism $g: (G, j) \rightarrow (G', j')$ such that $g \circ f_\lambda = f'_\lambda$. If $H = \{e\}$, then the amalgamated product is the same as the free product. Now f_λ is an injection. If we regard G_λ as a subgroup of G , then G is generated by the subgroups G_λ and $G_\lambda \cap G_\mu = j_\lambda(H) = j_\mu(H)$ ($\lambda \neq \mu$).

The notion of the amalgamated product is useful in constructing groups with interesting properties. For instance, we have a group whose nonidentity elements are all conjugate (B. H. Neumann and G. Higman), and a group generated by a finite number of elements such that its homomorphic image ($\neq \{e\}$) is always an infinite group (Higman) so that we have an infinite simple group generated by a finite number of elements.

N. Extensions

Let N and F be groups. A group G is called an **extension** of F by N if G has a normal subgroup \bar{N} isomorphic to N and $G/\bar{N} \cong F$. The problem of finding all extensions was solved by Schreier (*Monatsh. Math. Phys.*, 34 (1926); *Abh. Math. Sem. Univ. Hamburg*, 4 (1928)). Suppose that (1) to each $\sigma \in F$ there corresponds an automorphism s_σ of N ; (2) there exist elements $c_{\sigma, \tau}$ ($\sigma, \tau \in F$) of N such that $s_\sigma(s_\tau(a)) = c_{\sigma, \tau}(s_{\sigma\tau}(a))c_{\sigma, \tau}^{-1}$ ($a \in N$); and (3) $c_{\sigma, \tau}c_{\sigma\tau, \rho} = s_\sigma(c_{\tau, \rho})c_{\sigma, \tau\rho}$. Then the set G of all symbols as_σ ($a \in N, \sigma \in F$) is an extension of F by N if we define multiplication by $as_\sigma \cdot bs_\tau = (as_\sigma(b)c_{\sigma, \tau})s_{\sigma\tau}$. In fact, the set of all elements $\bar{a} = ac_{1,1}^{-1}s_1$ ($a \in N$) is a normal subgroup \bar{N} of G such that $G/\bar{N} \cong F$. Any extension can be obtained in this way. A system $(s_\sigma, c_{\sigma, \tau})$ satisfying (1), (2), and (3) above is called a **factor set** belonging to F . Two factor sets $(s_\sigma, c_{\sigma, \tau})$ and $(t_\sigma, d_{\sigma, \tau})$ are said to be **associated** if there exist elements a_σ ($\sigma \in F$) of N such that $t_\sigma(a) = s_\sigma(a_\sigma a a_\sigma^{-1})$ and $d_{\sigma, \tau} = a_\sigma(s_\sigma(a_\tau))c_{\sigma, \tau}a_{\sigma\tau}^{-1}$. In this case, two extensions determined by these factor sets are isomorphic. If $(s_\sigma, c_{\sigma, \tau})$ is associated with $(t_\sigma, d_{\sigma, \tau})$ ($d_{\sigma, \tau} = 1$ for any $\sigma, \tau \in F$), then we say that the corresponding extension is a **split extension**. In this case, the extension G contains a subgroup \bar{F} isomorphic to F , and $G = \bar{F}\bar{N}$, $\bar{F} \cap \bar{N} = \{e\}$. We call such an extension a **semidirect product** of N and F .

If N is Abelian, then condition (2) is simply $s_\sigma(s_\tau(a)) = s_{\sigma\tau}(a)$, since the only inner automorphism of N is the identity mapping. The con-

ditions of associated factor sets and split extension are also simplified. If N is contained in the center of G , then G is called a **central extension** of N .

O. Transfers

Let H be a subgroup of finite index in G and g_i ($i = 1, \dots, h$) be representatives of the right cosets of H . For $b \in Hg_i$ we write $g_i = \bar{b}$. Then for $x \in G$ an element $\bar{x} = H' \prod_{i=1}^h g_i x (g_i x)^{-1}$ of H/H' is determined uniquely (independent of the choice of representatives), where H' is the commutator subgroup of H . The correspondence $G'x \rightarrow \bar{x}$ yields a homomorphism of G/G' to H/H' , which is called the **transfer** from G/G' to H/H' .

P. Generalizations

The concept of group can be generalized in several ways. A set S in which a multiplication $(a, b) \rightarrow ab$ satisfying $(ab)c = a(bc)$ (the associative law) is defined is called a **semigroup**. If S is a commutative semigroup in which $ax = bx$ implies $a = b$ (the **cancellation law**), then S can be embedded in a group G so that the multiplication in S is preserved in G and any $x \in G$ is the quotient of two elements of S : $x = a^{-1}b = ba^{-1}$ ($a, b \in S$). Such a group G is determined uniquely by S . We call it the **group of quotients** of S .

The notion of semigroup is obtained by taking only associativity from the group axioms. On the other hand, if Q is a set with a law of composition $(a, b) \rightarrow ab$ which is not necessarily associative but satisfies the condition that any two among a, b, c in the equation $ab = c$ determine the third uniquely, then Q is called a **quasigroup**. A quasigroup with an identity element e such that $ea = ae = a$ for every element a is called a **loop**. For loops, we have an analog of the structure theory of groups (R. H. Bruck, *Trans. Amer. Math. Soc.*, 60 (1946)).

If we give up the possibility of forming products for all pairs of elements or the uniqueness of the product in the axioms for groups, then we have the following generalizations of groups. A set M with multiplication under which to any elements $a, b \in M$ there corresponds a nonempty subset ab of M is called a **hypergroupoid**. Moreover, if the associative law $(ab)c = a(bc)$ holds and for any elements $a, b \in M$ there exist $x, y \in M$ such that $b \in xa, b \in ay$, then M is called a **hypergroup**.

A set M is called a **mixed group** if (1) M can be partitioned into disjoint subsets M_0, M_1, M_2, \dots ; (2) for $a \in M_0, b \in M_i$ ($i = 0, 1, 2, \dots$),

elements $ab, a \setminus b$ of M_i are defined such that $a(a \setminus b) = b$; (3) for $b, c \in M_i$, an element b/c of M_0 is defined such that $(b/c) \cdot c = b$; and (4) the associative law $(ab)c = a(bc)$ ($a, b \in M_0, c \in M$) holds (A. Loewy, 1927).

A set M is called a **groupoid** if (1) M can be partitioned into disjoint subsets M_{ij} ($i, j = 1, 2, \dots$); (2) for $a \in M_{ij}$ and $b \in M_{jk}$, an element $ab \in M_{ik}$ is defined; (3) for $a \in M_{ij}$ and $b \in M_{ik}$, an element $a \setminus b \in M_{jk}$ is defined such that $a(a \setminus b) = b$; (4) for $a \in M_{ij}$ and $b \in M_{kj}$, an element $a/b \in M_{ik}$ is defined such that $(a/b)b = a$; and (5) for $a \in M_{ij}$, $b \in M_{jk}$, and $c \in M_{kl}$, the associative law $a(bc) = (ab)c$ holds (H. Brandt, 1926). These generalized concepts also have some practical applications (\rightarrow 2 Abelian Groups; 13 Algebraic Groups; 60 Classical Groups; 69 Compact Groups; 92 Crystallographic Groups; 122 Discontinuous Groups; 151 Finite Groups; 161 Free Groups; 243 Lattices; 249 Lie Groups; 277 Modules; 362 Representations; 422 Topological Abelian Groups; 423 Topological Groups; 437 Unitary Representations).

Q. History

The concept of the group was first introduced in the early 19th century, but its rudiments can be found in antiquity; in fact, it was virtually contained in the concept of motion or transformation used in ancient geometry. From the time it took explicit form in the late 19th century, it has played a fundamental role in all fields of mathematics.

In their study of algebraic equations in the late 18th century, J. L. †Lagrange, A. T. Vandermonde, and P. Ruffini saw the importance of the group of permutations of roots; using this idea N. H. Abel showed that a general equation of degree ≥ 5 cannot be solved algebraically. A. L. †Cauchy studied the group of permutations of roots for its own interest, but a complete description of the relationship between groups and algebraic equations was first given by E. †Galois. C. Jordan developed a detailed exposition of the theory given by Abel and Galois in his *Traité des substitutions* (1870) [1]. Up to that time, a group meant a permutation group; the axiomatic definition of a group was given by A. Cayley (1854) and L. Kronecker (1870). F. †Klein emphasized the significance of group theory in geometry in his †Erlangen program (1872), and M. S. †Lie developed the theory of †Lie groups in the 1880s. In 1897, W. Burnside published his *Theory of groups* [3], whose second edition (1911) is one of the classics in group theory and is still valuable. Since 1896, G. Frobenius [2] and others have developed the theory of

representation of groups by matrices (\rightarrow 362 Representations). By that time, the theory of finite groups had acquired all its essential features. Among the branches of abstract algebra, the theory of groups was the first to develop; it led to the progress of abstract algebra in the 1930s. Since the latter half of that decade, the theory of finite groups has been developed further; there has been increased interest in the theory, and many significant results have been obtained, especially since 1955 (\rightarrow 151 Finite Groups).

References

- [1] C. Jordan, *Traité des substitutions et des équations algébriques*, Gauthier-Villars, 1870 (Blanchard, 1957).
- [2] G. Frobenius, *Gesammelte Abhandlungen I–III*, Springer, 1968.
- [3] W. Burnside, *Theory of groups of finite order*, Cambridge Univ. Press, second edition, 1911.
- [4] H. Weyl, *The classical groups*, Princeton Univ. Press, revised edition, 1946.
- [5] B. L. van der Waerden, *Algebra*, Springer, I, seventh edition, 1966; II, fifth edition, 1967.
- [6] B. L. van der Waerden, *Gruppen von linearen Transformationen*, Erg. Math., Springer, 1935 (Chelsea, 1948).
- [7] A. G. Kurosh, *The theory of groups I, II*, Chelsea, 1960. (Original in Russian, 1953.)
- [8] M. Suzuki, *Structure of a group and the structure of its lattice of subgroups*, Erg. Math., Springer, 1956.
- [9] M. Hall, *The theory of groups*, Macmillan, 1959.
- [10] H. Zassenhaus, *Lehrbuch der Gruppentheorie I*, Teubner, 1937; English translation, *The theory of groups*, Chelsea, 1958.
- [11] A. Speiser, *Die Theorie der Gruppen von endlicher Ordnung*, Springer, third edition, 1937.
- [12] W. Specht, *Gruppentheorie*, Springer, 1956.
- [13] R. H. Bruck, *A survey of binary systems*, Erg. Math., Springer, 1958.
- [14] A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups*, Amer. Math. Soc. Math. Surveys, I, 1961; II, 1967.
- [15] E. S. Liapin, *Semigroups*, Amer. Math. Soc. Transl. of Math. Monographs, 1963.

191 (VII.8) G-Structures

Differential geometry studies differentiable manifolds and geometric objects or structures

on them. Among the geometric structures, the Riemannian and complex structures, with their contacts with other fields of mathematics and with their richness in results, occupy a central position in differential geometry. Perhaps it is impossible to say precisely what a differential geometric structure is or should be. However, the notion of **G-structure** allows a unified description of many of the interesting known geometric structures, such as Riemannian and complex structures.

A. The Notion of G-Structures

Let M be an m -dimensional C^∞ -manifold, and let $\tau: T(M) \rightarrow M$ be its \dagger tangent bundle. For $x \in M$, $T_x(M) = \tau^{-1}(x)$ is the vector space of tangent vectors at x . Let $\pi: F(M) \rightarrow M$ denote the \dagger frame bundle of M . It is a $GL(m; \mathbf{R})$ -principal bundle on M and $F_x(M) = \pi^{-1}(x)$ is the set of linear isomorphisms (called **frames** at x) of \mathbf{R}^m onto $T_x(M)$. Here $GL(m; \mathbf{R})$ is the general linear group of \mathbf{R}^m . Hence $F_x(M)$ can be naturally identified with the set of ordered bases of $T_x(M)$. Write $\theta = (\theta^1, \dots, \theta^m)$ for the **canonical form** of $F(M)$, which is the \mathbf{R}^m -valued 1-form on $F(M)$ defined by $\theta(v) = p^{-1}(\pi_*(v))$ for any $v \in T_p(F(M))$. Given a diffeomorphism $f: M \cong N$ of M onto another C^∞ -manifold N , we can naturally define the bundle isomorphism $f^{(1)}: F(M) \cong F(N)$ by $f^{(1)}(p) = f_* \circ p$.

Let G be a \dagger Lie subgroup of $GL(m; \mathbf{R})$. A principal G -subbundle $\pi_p: P \rightarrow M$ of the frame bundle $\pi: F(M) \rightarrow M$ is called a **G-structure** over M . Thus P is a regular submanifold of $F(M)$ satisfying the following three conditions:

(A1) $\pi(P) = M$.

(A2) for $p \in P$ and $\sigma \in GL(m; \mathbf{R})$, $p\sigma \in P$ if and only if $\sigma \in G$.

(A3) for any $x \in M$, there exist an open neighborhood U of x and a C^∞ -mapping $s: U \rightarrow P$ such that $\pi(s(y)) = y$ for any $y \in U$. Conversely, if a regular submanifold P of $F(M)$ satisfies the above three conditions (A1), (A2), and (A3), then $\pi_p = \pi|_P: P \rightarrow M$ is a G -structure over M . When G is closed, then condition (A3) is automatically satisfied. The restriction of θ onto P is called the **canonical form** of P and is also denoted by θ . For an open subset U of M , the restriction $P|_U = \pi_p^{-1}(U)$ is a G -structure over U .

Let $\pi_p: P \rightarrow M$ and $\pi_q: Q \rightarrow N$ be G -structures over M and N , respectively. A diffeomorphism $f: M \cong N$ is called a **G-isomorphism** of P onto Q if $f^{(1)}(P) = Q$. When such an f exists, we say that the G -structures P and Q are **equivalent**. A G -isomorphism of P onto itself is called a G -automorphism of P . Write $\text{Aut}(M, P)$ for the group of G -automorphisms of P .

Let \mathfrak{g} be the Lie algebra of G . Since G acts on P from the right, we have the natural Lie algebra homomorphism $\iota: \mathfrak{g} \rightarrow \Gamma(P, T(P))$. Here $\Gamma(P, T(P))$ is the Lie algebra of C^∞ -vector fields on P . Actually ι is injective, and we write A^* for $\iota(A)$ ($A \in \mathfrak{g}$). We call A^* the **fundamental vector field** corresponding to A .

For $x \in \mathbf{R}^m$, define $p_0(x) \in F_x(\mathbf{R}^m)$ by $p_0(x)((v^j)) = \sum_{j=1}^m v^j (\partial/\partial x^j)_x$, where x^1, \dots, x^m are the natural coordinates of \mathbf{R}^m . Any open subset D of \mathbf{R}^m has the natural G -structure $P(D, G)$ defined by $P(D, G) = \{p_0(x) \cdot \sigma \in F_x(\mathbf{R}^m); x \in D, \sigma \in G\}$. We say that a G -structure $\pi_p: P \rightarrow M$ over M is **integrable** if for each $x \in M$, there exists a local coordinate system (U, φ, D) around x such that φ is a G -isomorphism of $P|_U$ onto $P(D, G)$. Equivalently, if we set $\varphi = (x^1, \dots, x^m)$, then the frame $\{(\partial/\partial x^1)_y, \dots, (\partial/\partial x^m)_y\}$ belongs to P for each $y \in U$.

B. Examples

The following examples (B1)–(B7) show that some classical geometric structures can be treated uniformly from the viewpoint of G -structures.

(B1) An **absolute parallelism** on M is, by definition, a system $\{X_1, \dots, X_m\}$ of C^∞ -vector fields on M such that at each point $x \in M$ $\{X_1(x), \dots, X_m(x)\}$ is a basis of $T(M)_x$. Clearly this is equivalent to giving an $\{e\}$ -structure $\pi_p: P \rightarrow M$, where e stands for the identity matrix of $GL(m; \mathbf{R})$. Then P is integrable if and only if $[X_i, X_j] = 0$, $1 \leq i, j \leq m$. In this case, $\text{Aut}(M, P)$ is a finite-dimensional \dagger Lie transformation group on M .

(B2) Let E be a k -dimensional \dagger distribution, namely, E is a k -dimensional vector subbundle of $T(M)$. Write $\mathbf{R}^m = \mathbf{R}^k \oplus \mathbf{R}^{m-k}$. Set $GL(k, m; \mathbf{R}) = \{\sigma \in GL(m; \mathbf{R}); \sigma(\mathbf{R}^k) \subset \mathbf{R}^k\}$ and $P = \{p \in F(M); p(\mathbf{R}^k) \subset E\}$. Then P is a $GL(k, m; \mathbf{R})$ -structure over M . This gives a bijective correspondence between the k -dimensional distributions on M and the $GL(k, m; \mathbf{R})$ -structures on M . Moreover, P is integrable if and only if E is \dagger involutive (Frobenius theorem).

(B3) Let g be a Riemannian metric on M . Set $O(m) = \{\sigma \in GL(m; \mathbf{R}); {}^t\sigma\sigma = e\}$. Define $P(g)$ by $P(g) = \{(v_1, \dots, v_m) \in F(M); g(v_i, v_j) = \delta_{ij}\}$. Then $P(g)$ is an $O(m)$ -structure on M . This gives a bijective correspondence between the Riemannian metrics on M and the $O(m)$ -structures over M . Then $\text{Aut}(M, P(g))$ is the group of \dagger isometries of g , and a finite-dimensional Lie transformation group on M . Moreover, $P(g)$ is integrable if and only if g is flat in the sense that the Riemannian curvature tensor of g is zero.

(B4) Two Riemannian metrics g_1 and g_2 are

called **conformally equivalent** if there exists a C^∞ -function $\rho > 0$ with $g_1 = \rho g_2$. This is an equivalence relation among the Riemannian metrics on M . An equivalence class $\{g\}$ is called a **conformal structure** on M . Set $CO(m) = \{aA; A \in O(m), a \in \mathbf{R} - \{0\}\}$. Define $P(\{g\})$ by $P(\{g\}) = \{(v_1, \dots, v_m) \in F(M); g(v_i, v_j) = a\delta_{ij}, a \in \mathbf{R} - \{0\}\}$. Then $P(\{g\})$ is a $CO(m)$ -structure on M . This gives a bijective correspondence between the conformal structures on M and the $CO(m)$ -structures over M . Then $\text{Aut}(M, P(\{g\}))$ is the group of conformal transformations of $\{g\}$ and is a finite-dimensional Lie transformation group. Moreover $P(\{g\})$ is integrable if and only if the conformal structure $\{g\}$ is **conformally flat** in the sense that at each point $x \in M$, there exists a local coordinate system (x^1, \dots, x^m) around x such that $g(\partial/\partial x^i, \partial/\partial x^j) = \rho \delta_{ij}$, where ρ is a positive function.

(B5) Assume that m is even, say $m = 2l$. By an **almost symplectic structure** on M , we mean a differential 2-form Ω of maximal rank, i.e., $\Omega \wedge \dots \wedge \Omega$ (l times) never vanishes. Let A be the standard skew-symmetric bilinear form on \mathbf{R}^{2l} defined by $A((u^i), (v^j)) = (u^1 v^2 - u^2 v^1) + \dots + (u^{2l-1} v^{2l} - u^{2l} v^{2l-1})$. The symplectic group $Sp(l; \mathbf{R})$ is defined by $Sp(l; \mathbf{R}) = \{\sigma \in GL(2l; \mathbf{R}); A(\sigma u, \sigma v) = A(u, v), u, v \in \mathbf{R}^{2l}\}$. Set $P(\Omega) = \{p \in F(M); p^* A = \Omega_x, x = \pi_P(p)\}$. Then $P(\Omega)$ is an $Sp(l; \mathbf{R})$ -structure over M . This gives a bijective correspondence between the almost symplectic structures on M and the $Sp(l; \mathbf{R})$ -structures over M . Then $\text{Aut}(M, P(\Omega))$ is the group of symplectic transformations of Ω , which is never a finite-dimensional Lie transformation group. Moreover $P(\Omega)$ is integrable if and only if Ω is a symplectic structure, i.e., $d\Omega = 0$.

(B6) Assume that m is even, say $m = 2l$. We identify \mathbf{C}^l with \mathbf{R}^{2l} as a real vector space. Then $GL(m; \mathbf{C})$ is a Lie subgroup of $GL(2l; \mathbf{R})$. Let J be an almost complex structure on M . Set $P(J) = \{p \in F(M); Jp(u) = p(iu), u \in \mathbf{R}^{2l} = \mathbf{C}^l\}$. Then $P(J)$ is a $GL(l; \mathbf{C})$ -structure on M . This gives a bijective correspondence between the almost complex structures on M and the $GL(l; \mathbf{C})$ -structures over M . Then $\text{Aut}(M, P(J))$ is the group of J -analytic transformations on M . Furthermore, $\text{Aut}(M, P(J))$ is a finite-dimensional Lie transformation group when M is compact. Moreover $P(J)$ is integrable if and only if J is a complex structure.

(B7) It is easy to see that there is a bijective correspondence between the $SL(m; \mathbf{R})$ -structures over M and the set of volume elements on M . Any $SL(m; \mathbf{R})$ -structure over M is always integrable and $\text{Aut}(M, P)$ is the group of diffeomorphisms preserving Ω , which is never a finite-dimensional Lie transformation group.

C. Structure Functions

Let G be a Lie subgroup of $GL(m; \mathbf{R})$ and \mathfrak{g} the Lie algebra of G . For convenience, we write V for \mathbf{R}^m and V^* for the dual space of V . Then $V \otimes \wedge^2 V^*$ can be considered as the space of skew-symmetric bilinear mappings from $V \times V$ into V , and $\mathfrak{g} \otimes V^*$ can be identified with the space of linear mappings from V into \mathfrak{g} . Define a linear mapping $\partial: \mathfrak{g} \otimes V^* \rightarrow V \otimes \wedge^2 V^*$ by

$$(\partial T)(u, v) = -T(u)v + T(v)u$$

$$\text{for } T \in \mathfrak{g} \otimes V^*, u, v \in V.$$

Put

$$H^{2,1}(\mathfrak{g}) = V \otimes \wedge^2 V^* / \partial(\mathfrak{g} \otimes V^*).$$

Let $\pi_P: P \rightarrow M$ be a G -structure over M and θ the canonical form of P . Take any frame p in P . An m -dimensional linear subspace H of $T_p(P)$ is said to be a **horizontal subspace** of P at p if $\theta_p: H \rightarrow V (= \mathbf{R}^m)$ is a linear isomorphism. Then we write f_H for $\theta_p^{-1}: V \rightarrow H$. Write $\mathfrak{H}(T_p(P))$ for the set of horizontal subspaces of P at p . For $H \in \mathfrak{H}(T_p(P))$, define $c(p, H) \in V \otimes \wedge^2 V^*$ by

$$c(p, H)(u, v) = d\theta(f_H(u), f_H(v)) \quad \text{for } u, v \in V.$$

Then we have

$$c(p, H_1) - c(p, H_2) \in \partial(\mathfrak{g} \otimes V^*)$$

$$\text{for } H_1, H_2 \in \mathfrak{H}(T_p(P)).$$

Therefore we have a well-defined element $c(p) \in H^{2,1}(\mathfrak{g})$ by $c(p) = \{c(p, H)\}$. We call the mapping $c: P \rightarrow H^{2,1}(\mathfrak{g})$ the **structure function** of the G -structure $\pi_P: P \rightarrow M$.

Let $Q \rightarrow N$ be another G -structure over N and $f: M \simeq N$ be a G -isomorphism of P onto Q . Then we have $c(f^{(1)}(p)) = c(p)$ for $p \in P$. In particular, it is easily seen that if P is integrable, then $c \equiv 0$. However, the converse is not true in general.

D. Prolongation of Linear Lie Algebras and Groups

Let \mathbf{K} denote either \mathbf{R} or \mathbf{C} . Let \mathfrak{g} be a Lie subalgebra of $\mathfrak{gl}(m; \mathbf{K})$. For $k = 0, 1, 2, \dots$, let \mathfrak{g}_k be the space of \mathbf{K} -symmetric multilinear mappings

$$t: \underbrace{\mathbf{K}^m \times \dots \times \mathbf{K}^m}_{(k+1) \text{ times}} \rightarrow \mathbf{K}^m$$

such that for each fixed $u_1, \dots, u_k \in \mathbf{K}^m$, the linear transformation

$$v \in \mathbf{K}^m \mapsto t(v, u_1, \dots, u_k) \in \mathbf{K}^m$$

belongs to \mathfrak{g} . In particular, $\mathfrak{g}_0 = \mathfrak{g}$. We call \mathfrak{g}_k

the k th prolongation of \mathfrak{g} . If $\mathfrak{g}_k = 0$, then $\mathfrak{g}_{k+1} = \mathfrak{g}_{k+2} = \dots = 0$. The first integer k such that $\mathfrak{g}_k = 0$ is called the **order** of \mathfrak{g} . If $\mathfrak{g}_k \neq 0$ for all k , then \mathfrak{g} is said to be of **infinite type**. When \mathfrak{g} contains a matrix of rank 1 as an element, then \mathfrak{g} is of infinite type. When $\mathbf{K} = \mathbf{C}$, the converse is also true [3]. However, when $\mathbf{K} = \mathbf{R}$, the converse is not true in general. In fact, if we consider $\mathfrak{gl}(l; \mathbf{C})$ as a real Lie subalgebra of $\mathfrak{gl}(2l; \mathbf{R})$ (example (B6)), then $\mathfrak{gl}(l; \mathbf{C})$ is of infinite type, having no matrix of rank 1 in it. In general, a Lie subalgebra \mathfrak{g} of $\mathfrak{gl}(m; \mathbf{R})$ is said to be of **elliptic type** if \mathfrak{g} contains no matrix of rank 1 as an element. We consider \mathfrak{g}_1 as an Abelian Lie subalgebra of $\mathfrak{gl}(\mathbf{K}^m \oplus \mathfrak{g})$ by the correspondence $t \in \mathfrak{g}_1 \mapsto \bar{t} \in \mathfrak{gl}(\mathbf{K}^m \oplus \mathfrak{g})$, where \bar{t} is defined by

$$\begin{aligned}\bar{t}(u) &= t(\cdot, u) \quad \text{for } u \in \mathbf{K}^m, \\ \bar{t}(A) &= 0 \quad \text{for } A \in \mathfrak{g}.\end{aligned}$$

More generally, we consider \mathfrak{g}_{k+1} to be an Abelian Lie subalgebra of $\mathfrak{gl}(\mathbf{K}^m \oplus \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_k)$ by virtue of the correspondence $t \in \mathfrak{g}_{k+1} \mapsto \bar{t} \in \mathfrak{gl}(\mathbf{K}^m \oplus \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_k)$, where \bar{t} is defined by

$$\begin{aligned}\bar{t}(u) &= t(\cdot, \dots, \cdot, u) \in \mathfrak{g}_k \quad \text{for } u \in \mathbf{K}^m, \\ \bar{t}(A) &= 0 \quad \text{for } A \in \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_k.\end{aligned}$$

Identifying $\mathfrak{gl}(\mathbf{K}^m \oplus \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_{k-1}) \oplus \mathfrak{g}_k$ with $\mathfrak{gl}(\mathbf{K}^m \oplus \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_k)$, we have the important identity

$$(D1) \quad (\mathfrak{g}_k)_1 = \mathfrak{g}_{k+1}, \quad k=0, 1, 2, \dots$$

Let G be a Lie subgroup of $GL(m; \mathbf{R})$ and \mathfrak{g} the Lie subalgebra of $\mathfrak{gl}(m; \mathbf{R})$ corresponding to G . For $k=1, 2, \dots$, let G_k be the connected Abelian Lie subgroup of $GL(\mathbf{R}^m \oplus \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_{k-1})$ corresponding to the Abelian Lie subalgebra \mathfrak{g}_k ; namely, G_k consists of the linear mappings $\sigma(\bar{t})$ ($t \in \mathfrak{g}_k$) defined by

$$\begin{aligned}\sigma(\bar{t})(u) &= u + t(\cdot, \dots, \cdot, u) \quad \text{for } u \in \mathbf{R}^m, \\ \sigma(\bar{t})(A) &= A \quad \text{for } A \in \mathfrak{g}_0 \oplus \dots \oplus \mathfrak{g}_{k-1}.\end{aligned}$$

We call G_k the k th prolongation of G . We say that G is of finite (resp. infinite) type if \mathfrak{g} is of finite (resp. infinite) type. We say that G is of elliptic type if \mathfrak{g} is. From (D1), we have

$$(D2) \quad (G_k)_1 = G_{k+1}.$$

We know all irreducible Lie subalgebras \mathfrak{g} of $\mathfrak{gl}(m; \mathbf{K})$ with order ≥ 2 [3, 6].

(D3) An irreducible Lie subalgebra \mathfrak{g} of $\mathfrak{gl}(m; \mathbf{C})$ is of infinite type if and only if \mathfrak{g} is one of

$$\mathfrak{gl}(m; \mathbf{C}), \quad \mathfrak{sl}(m; \mathbf{C}), \quad \mathfrak{csp}\left(\frac{m}{2}; \mathbf{C}\right), \quad \mathfrak{sp}\left(\frac{m}{2}; \mathbf{C}\right),$$

where $\mathfrak{csp}(m/2; \mathbf{C}) = \mathbf{C} \oplus \mathfrak{sp}(m/2, \mathbf{C})$.

(D4) An irreducible Lie algebra \mathfrak{g} of $\mathfrak{gl}(m; \mathbf{R})$ is of infinite type if and only if \mathfrak{g} is one of

$$\begin{aligned}\mathfrak{gl}\left(\frac{m}{2}; \mathbf{C}\right), \quad \mathfrak{sl}\left(\frac{m}{2}; \mathbf{C}\right), \quad \mathfrak{csp}\left(\frac{m}{4}; \mathbf{C}\right), \\ \mathfrak{sp}\left(\frac{m}{4}; \mathbf{C}\right), \\ \mathfrak{gl}(m; \mathbf{R}), \quad \mathfrak{sl}(m; \mathbf{R}), \quad \mathfrak{csp}\left(\frac{m}{2}; \mathbf{R}\right), \quad \mathfrak{sp}\left(\frac{m}{2}; \mathbf{R}\right),\end{aligned}$$

where $\mathfrak{csp}(m/2; \mathbf{R}) = \mathbf{R} \oplus \mathfrak{sp}(m/2; \mathbf{R})$.

(D5) Let S be an m -dimensional irreducible Hermitian symmetric space of compact type, which is different from the complex projective space. Let $L(S)$ be the identity connected component of the group of biholomorphic transformations of S . Take any point $o \in S$ and set $L_o(S) = \{\sigma \in L(S); \sigma(o) = o\}$. Let $\mathfrak{g}(S)$ be the linear isotropy Lie subalgebra of $\mathfrak{gl}(m; \mathbf{C})$ corresponding to $L(S)$. Then $\mathfrak{g}(S)$ is an irreducible Lie subalgebra of $\mathfrak{gl}(m; \mathbf{C})$ of order 2. Conversely, every irreducible Lie subalgebra \mathfrak{g} of $\mathfrak{gl}(m; \mathbf{C})$ of finite type with order ≥ 2 is equal to $\mathfrak{g}(S)$ as given above.

(D6) Let S be an m -dimensional irreducible symmetric space of compact type. Assume that S admits a finite-dimensional connected Lie transformation group $L(S)$ which strictly contains the connected component of the group of isometries of S . Take any point $o \in S$, and set $L_o(S) = \{\sigma \in L(S); \sigma(o) = o\}$. Let $\mathfrak{g}(S)$ be the linear isotropy Lie subalgebra of $\mathfrak{gl}(m; \mathbf{R})$ corresponding to $L_o(S)$. If $(S, L(S)) \neq (\text{sphere, projective transformation group})$ or $(S, L(S)) \neq (\text{complex projective space, complex projective transformation group})$, then $\mathfrak{g}(S)$ is an irreducible Lie subalgebra of $\mathfrak{gl}(m; \mathbf{R})$ of order 2. Conversely, every irreducible Lie subalgebra \mathfrak{g} of $\mathfrak{gl}(m; \mathbf{R})$ of finite type with order ≥ 2 is equal to $\mathfrak{g}(S)$ as given above. For example, if we take an m -dimensional sphere as S and the group of conformal transformations of S as $L(S)$, then $\mathfrak{g}(S) = \mathfrak{co}(m)$.

E. Prolongation of G-Structures

Let G be a Lie subgroup of $GL(m; \mathbf{R})$ and let \mathfrak{g} be the Lie subalgebra of $\mathfrak{gl}(m; \mathbf{R})$ corresponding to G . We choose once and for all a linear subspace C of $V \otimes \wedge^2 V^*$ such that

$$V \otimes \wedge^2 V^* = \partial(\mathfrak{g} \otimes V^*) \oplus C.$$

In general there is no natural way of choosing such a C .

Let $\pi_p: P \rightarrow M$ be a G -structure on M . For each horizontal subspace H of P at p (i.e., $H \in \Theta(T_p(P))$), we define the frame $(p, H) \in F_p(P)$

to be a linear isomorphism $(p, H): \mathbf{R}^m \oplus \mathfrak{g} \rightarrow T_p(P)$ given by

$$(p, H)(v \oplus A) = f_H(v) \oplus A_p^* \quad \text{for } v \in \mathbf{R}^m, A \in \mathfrak{g},$$

where A^* is the fundamental vector field on P induced from $A \in \mathfrak{g}$. Let $P_1 = \{(p, H) \in F(P); H \in \mathcal{H}(T_p(P)), c(p, H) \in C\}$. Then $P_1 \rightarrow P$ is a G_1 -structure over P . We call $P_1 \rightarrow P$ the **first prolongation** of P . The k th **prolongation** P_k of P is defined inductively by $P_k = (P_{k-1})_1$ = the first prolongation of P_{k-1} . From (D2), P_k is a G_k -structure over P_{k-1} .

Take any $f \in \text{Aut}(M, P)$. It can be proved that $f^{(1)} \in \text{Aut}(P, P_1)$. Since $P_k = (P_{k-1})_1$, we can define $f^{(k)} \in \text{Aut}(P_{k-1}, P_k)$ inductively by $f^{(k)} = (f^{(k-1)})^{(1)}$. By the correspondence $f \mapsto f^{(k)}$, we can consider $\text{Aut}(M, P)$ as a subgroup of $\text{Aut}(P_{k-1}, P_k)$.

If G is connected, then

$$(E1) \quad \text{Aut}(M, P) = \text{Aut}(P_{k-1}, P_k), \quad k = 1, 2, \dots$$

Let $Q \rightarrow N$ be another G -structure. By a similar argument to that above, we see that $P \rightarrow M$ and $Q \rightarrow N$ are equivalent if and only if $P_k \rightarrow P_{k-1}$ and $Q_k \rightarrow Q_{k-1}$ are equivalent. In [1], E. Cartan studied general equivalence problems of two G -structures. In that problem the above prolongation procedure plays an important role.

F. Automorphisms of G -Structures of Finite Type

S. Kobayashi proved the following:

(F1) Theorem [4]. Let $P \rightarrow M$ be an $\{e\}$ -structure over M . Then $\text{Aut}(M, P)$ is a finite-dimensional Lie transformation group on M such that $\dim \text{Aut}(M, P) \leq \dim M$. More precisely, for any point $x \in M$, the mapping $\sigma \in \text{Aut}(M, P) \mapsto \sigma(x) \in M$ is injective, and its image $\{\sigma(x); \sigma \in \text{Aut}(M, P)\}$ is a closed submanifold of M . The submanifold structure on this image makes $\text{Aut}(M, P)$ into a finite-dimensional Lie transformation group on M .

Now let G be a Lie subgroup of $GL(m; \mathbf{R})$ of finite type, say $G_k = \{e\}$. Then $P_k \rightarrow P_{k-1}$ is an $\{e\}$ -structure. From theorem (F1) above, it follows that $\text{Aut}(P_{k-1}, P_k)$ is a finite-dimensional Lie transformation group. As explained in Section E, $\text{Aut}(M, P)$ is a subgroup of $\text{Aut}(P_{k-1}, P_k)$. Clearly $\text{Aut}(M, P)$ is closed in $\text{Aut}(P_{k-1}, P_k)$. Since every closed subgroup of a Lie group is again a Lie group, we have the following:

(F2) Theorem. Let $P \rightarrow M$ be a G -structure over M . If G is of finite type, then $\text{Aut}(M, P)$ is a Lie transformation group of dimension $\leq \dim(\mathbf{R}^m \oplus \mathfrak{g} \oplus \mathfrak{g}_1 \oplus \dots \oplus \mathfrak{g}_{k-1})$.

G. Automorphisms of G -Structures of Elliptic Type

The following theorem of R. Palais allows us to prove that the automorphism groups of many general geometric structures are finite-dimensional Lie transformation groups.

(G1) Theorem [5]. Let L be a group of C^∞ -transformations of a C^∞ -manifold M . Let \mathcal{L} be the set of all vector fields X on M which generate global 1-parameter groups $\varphi_t = \exp tX$ of transformations of M such that $\varphi_t \in L$. If the set \mathcal{L} generates a finite-dimensional Lie algebra of vector fields on M , then L is a finite-dimensional Lie transformation group, and \mathcal{L} is the Lie algebra of L .

Let $P \rightarrow M$ be a G -structure over M . For any representation $\rho: G \rightarrow GL(V)$, we write $E(\rho)$ for the associated vector bundle of P with respect to ρ . Let $\rho_1: G \rightarrow GL(\mathfrak{gl}(m; \mathbf{R}))$ be the representation defined by $\rho_1(\sigma)A = \sigma A \sigma^{-1}$ for $\sigma \in G$, $A \in \mathfrak{gl}(m; \mathbf{R})$. Let $\rho_2: G \rightarrow GL(\mathfrak{g})$ be the representation defined by $\rho_2(\sigma)A = \sigma A \sigma^{-1}$ for $\sigma \in G$, $A \in \mathfrak{g}$. Then ρ_2 is a subrepresentation of ρ_1 , since $\rho_1(\sigma)A = \rho_2(\sigma)A$ for $\sigma \in G$, $A \in \mathfrak{g}$. We remark that $E(\rho_1) = \text{Hom}(T(M), T(M))$. We write $\mathfrak{g}(M)$ for $E(\rho_2)$. Then $\mathfrak{g}(M)$ is a vector subbundle of $\text{Hom}(T(M), T(M))$. We write F for the quotient vector bundle $\text{Hom}(T(M), T(M))/\mathfrak{g}(M)$. Let $\Phi: \text{Hom}(T(M), T(M)) \rightarrow F$ be the natural projection. We fix an affine connection ∇ on M which preserves P . We write T for the torsion tensor field of ∇ . For each vector field $X \in \Gamma(M, T(M))$, we define a C^∞ -section $\nabla X + T(X, \cdot)$ in $\Gamma(M, \text{Hom}(T(M), T(M)))$ by

$$u \in T(M)_x \mapsto \nabla_u X + T(X(x), u) \in T_x(M).$$

Then define a first-order linear differential operator

$$D: \Gamma(M, T(M)) \rightarrow \Gamma(M, F)$$

by $D(X) = \Phi(\nabla X + T(X, \cdot))$. It is easy to see that

$$(G2) \quad \mathcal{L}(M, P) \subset \ker D,$$

where $\mathcal{L}(M, P)$ is the Lie algebra of all vector fields $X \in \Gamma(M, T(M))$ which generate global 1-parameter groups $\varphi_t = \exp tX$ of transformations of M such that $\varphi_t \in \text{Aut}(M, P)$. Then we can show:

(G3) D is an elliptic operator if and only if G is of elliptic type.

Now suppose M is compact. Then from the standard fact on linear elliptic operators on compact manifolds, we know the dimension of $\ker D$ is finite if D is elliptic. Thus from (G2), (G3), and theorem (G1), we have the following:

(G4) Theorem [7]. Let $P \rightarrow M$ be a G -structure on an m -dimensional compact mani-

fold M . If G is of elliptic type, then $\text{Aut}(M, P)$ is a finite-dimensional Lie transformation group.

H. Local Equivalence Problem

Let $\pi_P: P \rightarrow M$ and $\pi_Q: Q \rightarrow N$ be two G -structures. We say that $P \rightarrow M$ and $Q \rightarrow N$ are **locally equivalent** if the following holds:

(H1) For arbitrarily given $(p, q) \in P \times Q$, there exists a local G -isomorphism $f: U \cong V$ of $P|U$ onto $Q|V$ such that $f^{(1)}(p) = q$, where U (resp. V) is an open neighborhood of $\pi_P(p)$ (resp. $\pi_Q(q)$).

From now on, we assume that the structure function c_P (resp. c_Q) of P (resp. Q) is constant. If such an f as in (H1) exists, we must have $c_Q(f^{(1)}(p')) = c_P(p')$ for $p' \in P|U$. Therefore c_P and c_Q must be the same constant. Cartan reduced the local equivalence problem between P and Q to that of certain differential systems on $P \times Q$. In fact, let $\theta_P = (\theta^1, \dots, \theta^m)$ and $\theta_Q = (\psi^1, \dots, \psi^m)$ be the canonical forms of P and Q respectively. Let $\alpha: P \times Q \rightarrow P$, $\beta: P \times Q \rightarrow Q$ be the canonical projections such that $\alpha(p, q) = p$, $\beta(p, q) = q$. Let Σ be the differential system on $P \times Q$ generated by $\{\alpha * \theta^1 - \beta * \psi^1, \dots, \alpha * \theta^m - \beta * \psi^m\}$. If X is the m -dimensional regular submanifold of $P \times Q$ given by the graph of $f^{(1)}$ in (H1), i.e., $X = \{(p', f^{(1)}(p')) \in P \times Q; p' \in P|U\}$, then X is an m -dimensional integral submanifold of Σ satisfying the following conditions:

$$(H2) \quad (p, q) \in X,$$

$$(H3) \quad \alpha * \theta^1, \dots, \alpha * \theta^m \text{ are linearly independent on } T_{(p,q)}(X).$$

Conversely any m -dimensional integral submanifold of Σ satisfying (H2) and (H3) is the graph of a local G -isomorphism f required in (H1). Therefore the local equivalence problem between P and Q is equivalent to the problem of finding an m -dimensional integral submanifold X of Σ satisfying (H2) and (H3) for any $(p, q) \in P \times Q$.

We say that G is **involutive** if there exist linear subspaces $0 = V_0 \subset V_1 \subset \dots \subset V_m = \mathbf{R}^m$ such that

$$(H4) \quad \dim V_k = k \quad (k = 0, \dots, m),$$

$$(H5) \quad \dim \mathfrak{g}_1 = \sum_{k=0}^m \dim \mathfrak{g}(V_k),$$

where \mathfrak{g} is the Lie algebra of G and $\mathfrak{g}(V_k) = \{A \in \mathfrak{g}; A(u) = 0 \text{ for } u \in V_k\}$. Now we have:

(H6) Proposition. Let $P \rightarrow M$ and $Q \rightarrow N$ be two G -structures such that their structure functions are constant and equal. The differential system Σ on $P \times Q$ defined as above is involutive at any m -dimensional integral

element E of Σ on which $\alpha * \theta^1, \dots, \alpha * \theta^m$ are linearly independent if and only if G is involutive.

Combining proposition (H6) above with the classical \dagger Cartan-Kähler theorem, we obtain the following theorems:

(H7) Theorem ([1], also see [8]). Let $P \rightarrow M$ and $Q \rightarrow N$ be two real analytic G -structures over M and N , respectively, such that their structure functions are constant and equal. Suppose G is involutive. Then $P \rightarrow M$ and $Q \rightarrow N$ are locally equivalent.

(H8) Theorem. Let $P \rightarrow M$ be a real analytic G -structure over M . Suppose G is involutive. Then P is integrable if and only if the structure function of P is zero everywhere.

Finally, we remark that for a general Lie subgroup $G \subset GL(m, \mathbf{R})$, G_k is involutive for $k \geq k_0$ [8].

I. Cartan-Kähler Theorem on Differential Systems

Let U be an open subset of \mathbf{R}^m . We write $A^k(U)$ for the space of real analytic differential k -forms on U . Put $A^*(U) = A^0(U) \oplus \dots \oplus A^m(U)$. Then $A^*(U)$ is a graded \mathbf{R} -algebra with respect to the usual exterior product. By a **differential system** on U we mean an ideal Σ of $A^*(U)$ such that

$$(I1) \quad \Sigma = A^0(U) \cap \Sigma \oplus A^1(U) \cap \Sigma \oplus \dots \oplus \Sigma \cap A^m(U),$$

$$(I2) \quad d(\Sigma) \subset \Sigma,$$

$$(I3) \quad \Sigma \text{ is finitely generated as an ideal.}$$

For convenience we write $\Sigma^{(k)} = \Sigma \cap A^k(U)$. When $\Sigma^{(0)} = \{0\}$, we call Σ a **restricted differential system**.

Let Σ be a differential system on $U \subset \mathbf{R}^m$. By a **k -dimensional integral manifold** of Σ , we mean a k -dimensional regular real analytic submanifold of U such that for any $\alpha \in \Sigma$, we have $\alpha|X \equiv 0$. In the above definition, it is sufficient to know that $\alpha|X \equiv 0$ for any $\alpha \in \Sigma^{(k)}$.

For $x \in U$, we write $G_k(T_x(U))$ for the set of all k -dimensional linear subspaces of $T_x(U)$. Set $G_k(T(U)) = \bigcup_{x \in U} G_k(T_x(U))$ (disjoint union); then $G_k(T(U))$ is naturally a real analytic manifold of dimension $m + m(m - k)$. Let Σ be a differential system on U . A k -plane $E \in G_k(T_x(U))$ is called a **k -dimensional integral element** of Σ at x if for any $\alpha \in \Sigma^{(k)}$, we have $\alpha|E = 0$. Thus X is a k -dimensional integral submanifold if and only if $T_x(X)$ is a k -dimensional integral element of Σ for each $x \in X$. We write $I_k(\Sigma)$ for the set of k -dimensional integral elements of Σ . In general $I_k(\Sigma)$ is a real analytic subset of $G_k(T(U))$. For any k -dimensional

integral element E at $x \in U$, we define the **polar space** $H(E)$ by

$$H(E) = \{u \in T_x(U); \alpha(u, v_1, \dots, v_k) = 0 \\ \text{for } \alpha \in \Sigma^{(k+1)}, v_1, \dots, v_k \in E\}.$$

Then $E \subset H(E)$. Moreover $F \in G_{k+1}(T_x(U))$ with $F \supset E$ is in $I_{k+1}(\Sigma)$ if and only if $F \subset H(E)$. Let E be a k -dimensional integral element of Σ . We choose $\varphi^1, \dots, \varphi^k \in A^1(U)$ such that $\varphi^1|E, \dots, \varphi^k|E$ are linearly independent. Define $\mathcal{U}(\varphi^1, \dots, \varphi^k)$ by

$$\mathcal{U}(\varphi^1, \dots, \varphi^k) = \{F \in G_k(T(U)); \varphi^1|F, \dots, \varphi^k|F \\ \text{linearly independent}\}.$$

Then $\mathcal{U}(\varphi^1, \dots, \varphi^k)$ is an open neighborhood of E in $G_k(T(U))$. For any element $F \in \mathcal{U}(\varphi^1, \dots, \varphi^k)$, define $F_1, \dots, F_k \in F$ by $\varphi^i(F_k) = \delta_k^i$. Thus $\{F_1, \dots, F_k\}$ is the dual basis of $\varphi^1|F, \dots, \varphi^k|F$. Any element $\alpha \in \Sigma^{(k)}$ defines a real analytic function α_* on $\mathcal{U}(\varphi^1, \dots, \varphi^k)$ by $\alpha_*(F) = \alpha(F_1, \dots, F_k)$. We call E **regular** if there exist $\alpha^1, \dots, \alpha^l \in \Sigma^{(k)}$ and an open neighborhood \mathcal{V} of E in $\mathcal{U}(\varphi^1, \dots, \varphi^k)$ such that

$$(I4) \quad I_k(\Sigma) \cap \mathcal{V} = \{F \in \mathcal{V}; \alpha_*^1(F) = \dots = \alpha_*^l(F) = 0\},$$

$$(I5) \quad d\alpha_*^1, \dots, d\alpha_*^l \text{ are linearly independent on } \mathcal{V},$$

$$(I6) \quad \dim H(F) \text{ is constant for any } F \in \mathcal{V}.$$

We say that $E \in I_k(\Sigma)$ is an **ordinary element** if E contains a $(k-1)$ -dimensional regular integral element.

The following is the well-known theorem due to Cartan and Kähler [7].

(17) Theorem. Let Σ be a restricted differential system on $U \subset \mathbf{R}^m$. Let X be a k -dimensional integral submanifold of Σ . Suppose that $T_x(X)$ is regular for a point $x \in X$ and that there exists a $(k+1)$ -dimensional integral element F at x with $F \supset T_x(X)$. Then there exists a $(k+1)$ -dimensional integral submanifold Y of Σ such that $x \in Y$, $T_x(Y) = F$, and $X \subset Y$ in a neighborhood of x .

We say that Σ is **involutive** at $E \in I_k(\Sigma)$ if there exist $0 = E^0 \subset E^1 \subset \dots \subset E^{k-1} \subset E$ such that E^j is a j -dimensional regular integral element of Σ ($j=0, 1, \dots, k-1$). Applying theorem (G7) inductively, we obtain:

Corollary. Let Σ be a restricted differential system on $U \subset \mathbf{R}^m$. If E is a k -dimensional involutive integral element of Σ at x , there exists a k -dimensional integral submanifold of Σ such that $x \in X$ and $T_x(X) = E$.

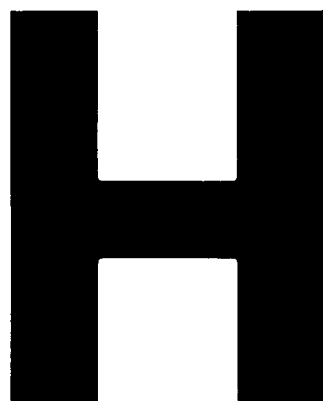
In view of this corollary, it is important to know when Σ is involutive at $E \in I_k(\Sigma)$.

(18) Lemma [9]. Let Σ be a restricted differential system on $U \subset \mathbf{R}^m$, and E be a k -dimensional integral element of Σ . Then Σ is

involutive at E if and only if there exist $0 = E^0 \subset E^1 \subset \dots \subset E^{k-1} \subset E$ such that $I_k(\Sigma)$ is a submanifold of $G_k(T(U))$ near E with dimension $m + k(m-k) - \sum_{j=0}^{k-1} t_j$, where $t_j = m - \dim H(E^j)$.

References

- [1] E. Cartan, Les problèmes d'équivalence, Oeuvres, pt. II, vol. 2, 1311–1334.
- [2] V. Guillemin and S. Sternberg, An algebraic model of transitive differential geometry, Bull. Amer. Math. Soc., 70 (1964), 16–47.
- [3] V. Guillemin, D. Quillen, and S. Sternberg, The classification of the irreducible complex algebras of infinite type, J. Analyse Math., 18 (1967), 107–112.
- [4] S. Kobayashi, Le groupe des transformations qui laissent invariant le parallélisme, Colloque de Topologie de Strasbourg, 1954.
- [5] S. Kobayashi, Transformation groups in differential geometry, Erg. Math., 70, Springer.
- [6] S. Kobayashi and T. Nagano, On filtered Lie algebras and geometric structures, J. Math. Mech. I, 13 (1964), 875–908; II, 14 (1965), 513–522; III, 14 (1965), 679–706; IV, 15 (1966), 163–175; V, 15 (1966), 315–328.
- [7] M. Spivak, A comprehensive introduction to differential geometry V, Publish or Perish Inc., 1975.
- [8] T. Ochiai, On the automorphism group of a G -structure, J. Math. Soc. Japan, 18 (1966), 189–193.
- [9] I. Singer and S. Sternberg, The infinite groups of Lie and Cartan: The transitive groups, J. Analyse Math., 15 (1965), 1–114.
- [10] N. Tanaka, On the equivalence problems associated with a certain class of homogeneous spaces, J. Math. Soc. Japan, 17 (1965), 103–139.



192 (X.24) Harmonic Analysis

A. Fourier Transforms

Let $f(x)$ be an element of the \dagger function space $L_1(-\infty, \infty)$ and t a real number. Then the integral

$$\hat{f}(t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(x) e^{-itx} dx \quad (1)$$

converges, and the function $\hat{f}(t)$ is continuous in $(-\infty, \infty)$. We call \hat{f} the \dagger Fourier transform of f . If $f(x) \in L_2(-\infty, \infty)$, then $f(x) \in L_1(-a, a)$ for any finite interval $(-a, a)$, and if we set

$$\hat{f}_a(t) = (2\pi)^{-1/2} \int_{-a}^a f(x) e^{-itx} dx,$$

then \hat{f}_a \dagger converges in the mean of order 2 as $a \rightarrow \infty$ to a function \hat{f} in L_2 . In this case, we define \hat{f} to be the \dagger Fourier transform of f ($\in L_2$). Furthermore, in this case, if we set

$$f_a(x) = (2\pi)^{-1/2} \int_{-a}^a \hat{f}(t) e^{itx} dt,$$

then $\lim_{a \rightarrow \infty} f_a(x) = f(x)$ (**Plancherel theorem**). Moreover, we have \dagger Parseval's identity: $\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\hat{f}(t)|^2 dt$ (\rightarrow 160 Fourier Transform).

Suppose that $f(x)$ is periodic with period 2π and $f \in L_2(-\pi, \pi)$. We set

$$a_n = (2\pi)^{-1} \int_{-\pi}^{\pi} f(x) e^{-inx} dx$$

(\dagger Fourier coefficients). The n th partial sum of the \dagger Fourier series $s_n(x) = \sum_{v=-n}^n a_v e^{ivx}$ converges in the mean of order 2 to $f(x)$, and Parseval's identity

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{n=-\infty}^{\infty} |a_n|^2$$

holds. On the other hand, if $\{a_n\}$ is a given sequence such that $\sum_{n=-\infty}^{\infty} |a_n|^2 < \infty$, then $\sum_{v=-n}^n a_v e^{ivx}$ converges in the mean of order 2 to a function $f(x)$, the Fourier coefficients of $f(x)$ are $\{a_n\}$, and Parseval's identity holds (\rightarrow 159 Fourier Series).

B. Bochner's Theorem and Herglotz's Theorem

A complex-valued function $f(x)$ defined on $(-\infty, \infty)$ is said to be of **positive type** (or **positive definite**) if $\sum_{j,k=1}^n f(x_j - x_k) \xi_j \bar{\xi}_k \geq 0$ for any finite number of reals x_1, x_2, \dots, x_n and complex numbers $\xi_1, \xi_2, \dots, \xi_n$. If $f(x)$ is measurable on $(-\infty, \infty)$ and of positive type, then there exists a \dagger monotone increasing real-

valued bounded function $\alpha(t)$ such that

$$f(x) = \int_{-\infty}^{\infty} e^{itx} d\alpha(t) \quad (2)$$

for almost all x . If $\alpha(-\infty) = 0$ and $\alpha(t)$ is right continuous, then $\alpha(t)$ is unique (**Bochner's theorem**). Conversely, if $\alpha(t)$ is nondecreasing and bounded and $f(x)$ is defined by (2), then $f(x)$ (called the **Fourier-Stieltjes transform** of $\alpha(t)$) is continuous and of positive type.

A sequence $\{a_n\}$ ($-\infty < n < \infty$) is said to be **positive definite** (or of **positive type**) if $\sum_{j,k=1}^n a_{j-k} \xi_j \bar{\xi}_k \geq 0$ for finitely many arbitrarily chosen complex numbers $\xi_1, \xi_2, \dots, \xi_n$. If $\{a_n\}$ is of positive type, then there exists a monotone increasing bounded function $\alpha(t)$ on $[-\pi, \pi]$ such that

$$a_n = \int_{-\pi}^{\pi} e^{int} d\alpha(t)$$

(**Herglotz's theorem**). Conversely, if $\alpha(t)$ is monotone increasing and bounded and a_n is defined by the above integral, then the sequence $\{a_n\}$ is of positive type.

C. Poisson's Summation Formula

If $f(x) \in L_1(-\infty, \infty)$ is of \dagger bounded variation and continuous and if $\hat{f}(t)$ is its Fourier transform, then we have

$$\sqrt{a} \sum_{k=-\infty}^{\infty} f(ak) = \sqrt{b} \sum_{k=-\infty}^{\infty} \hat{f}(bk),$$

where $ab = 2\pi$ ($a > 0$). This is called **Poisson's summation formula**.

D. Generalized Tauberian Theorems of Wiener

Suppose that we are given a function $f(x) \in L_1(-\infty, \infty)$ whose Fourier transform $\hat{f}(t)$ is never zero. Then the set of functions given by

$$h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy,$$

where $g \in L_1(-\infty, \infty)$, is dense in $L_1(-\infty, \infty)$. Hence we can deduce the \dagger generalized Tauberian theorem of Wiener: If the Fourier transform $\hat{k}_1(t)$ of $k_1(x) \in L_1(-\infty, \infty)$ does not vanish for any real t and

$$\lim_{x \rightarrow \infty} \int_{-\infty}^{\infty} k_1(x-y)f(y)dy = C \int_{-\infty}^{\infty} k_1(y)dy$$

for a function $f(x)$ that is bounded and measurable on $(-\infty, \infty)$, then for any $k_2 \in L_1(-\infty, \infty)$,

$$\lim_{x \rightarrow \infty} \int_{-\infty}^{\infty} k_2(x-y)f(y)dy = C \int_{-\infty}^{\infty} k_2(y)dy$$

(→ 160 Fourier Transform G). Hence we can deduce †Tauberian theorems of the Littlewood type [3].

E. Harmonic Analysis

Let $\alpha(\lambda)$ be a complex-valued function on $(-\infty, \infty)$ that is of bounded variation and right continuous. If we have the expression

$$f(t) = \int_{-\infty}^{\infty} e^{i\lambda t} d\alpha(\lambda), \quad (2')$$

then we say that the function $f(t)$ is represented by the superposition of harmonic oscillations $e^{i\lambda t}$. Conversely, when $f(t)$ is given, we have the problem of finding a function $\alpha(\lambda)$ as above such that $f(t)$ can be expressed in the form of (2'). When such a function $\alpha(\lambda)$ exists, it is also an important problem (in harmonic analysis) to find the amplitude $\alpha(\lambda) - \alpha(\lambda - 0)$ of the component of a proper oscillation. Concerning this problem, we have the following three theorems:

(I) A necessary and sufficient condition for $f(x)$ to be representable in the form (2') is

$$\sup_{1 \leq n} \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \left(\frac{\sin(t/n)}{t/n} \right)^2 f(t) e^{-i\lambda t} dt \right| d\lambda < \infty.$$

(II) For a function $f(t)$ expressed in the form (2'), we have: (i) for any λ_0 ,

$$\alpha(\lambda_0) - \alpha(\lambda_0 - 0) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t) e^{-i\lambda_0 t} dt,$$

and (ii) if $\alpha(\lambda)$ is continuous at $\lambda = \lambda_0 - \sigma$ and $\lambda = \lambda_0 + \sigma$ ($\sigma > 0$), then

$$\begin{aligned} & \alpha(\lambda_0 + \sigma) - \alpha(\lambda_0 - \sigma) \\ &= \lim_{T \rightarrow \infty} \sqrt{\frac{2}{\pi}} \int_{-T}^T \frac{\sin \sigma t}{t} f(t) e^{-i\lambda_0 t} dt. \end{aligned}$$

(III) In (2'), suppose that the discontinuity points of $\alpha(\lambda)$ are $\lambda_1, \lambda_2, \dots$, and set $a_n = \alpha(\lambda_n) - \alpha(\lambda_n - 0)$ ($n = 1, 2, \dots$); then

$$\lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S f(t+s) f(\bar{s}) ds = \sum_{n=1}^{\infty} |a_n|^2 e^{i\lambda_n t}.$$

F. The Paley-Wiener Theorem

In formula (1), if we change the variable from t to a complex variable $\zeta = t + i\sigma$, then we have

$$F(\zeta) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(x) e^{-i\zeta x} dx, \quad (3)$$

which is called the **Fourier-Laplace transform** of $f(x)$. In particular, if $f(x)$ has bounded †support, then $F(\zeta)$ is an †entire function. Concerning Fourier-Laplace transforms, Paley and Wiener proved the following theorems:

(I) A necessary and sufficient condition for an entire function $F(\zeta)$ to be the Fourier-Laplace transform of a †function of class C^∞ having its support in a finite interval $[-B, B]$ is that for any N , there exists a constant $C_N > 0$ such that $|F(\zeta)| \leq C_N (1 + |\zeta|)^{-N} e^{B|\sigma|}$ for all $\zeta = t + i\sigma$.

(II) If $g(t) \in L_2(0, \infty)$, then its one-sided †Laplace transform

$$f(z) = (2\pi)^{-1/2} \int_0^\infty g(t) e^{-tz} dt$$

satisfies: (i) $f(z)$ is holomorphic in the right half-plane $\operatorname{Re} z > 0$, and (ii)

$$\sup_{x>0} \int_{-\infty}^{\infty} |f(x+iy)|^2 dy < \infty.$$

Conversely, if $f(z) = f(x+iy)$ ($x > 0$) satisfies (i) and (ii), then the boundary function $f(iy) \in L_2(-\infty, \infty)$ exists and is such that

$$\lim_{x \downarrow 0} \int_{-\infty}^{\infty} |f(iy) - f(x+iy)|^2 dy = 0,$$

its Fourier transform in L_2

$$g(t) = \text{l.i.m.}_{N \rightarrow \infty} (2\pi)^{-1/2} \int_{-N}^N f(iy) e^{ity} dy$$

vanishes at almost all negative t , and $f(z)$ is the one-sided Laplace transform of $g(t)$.

G. Harmonic Analysis on Locally Compact Abelian Groups

The general theory of harmonic analysis on the real line was extended to a theory on locally compact Abelian groups by A. Weil, I. M. Gel'fand, D. A. Raikov, and others. The theory of normed rings was utilized for the development of the theory (→ 36 Banach Algebras). This theory is called **harmonic analysis on locally compact Abelian groups**, and it has been developed as described in the following sections.

H. Group Rings

Let $L_1 = L_1(G)$ be the set of all integrable functions with respect to a †Haar measure on a locally compact Abelian group G . If we define the norm and multiplication in $L_1(G)$ by

$$\|f\| = \int_G |f(x)| dx,$$

$$f \cdot g(x) = \int_G f(xy^{-1}) g(y) dy,$$

respectively, then L_1 has the structure of a commutative †Banach algebra. (We call $f \cdot g$ the **convolution** (or **composition product**) of f

and g .) If the topology of G is not † discrete, $L_1(G)$ does not have a unity for multiplication. Hence, adjoining a formal unity $\mathbf{1}$ to $L_1(G)$, we set $R = \{\alpha\mathbf{1} + f \mid \alpha \text{ is a complex number, } f \in L_1(G)\}$, and

$$\|\alpha\mathbf{1} + f\| = |\alpha| + \|f\|,$$

$$(\alpha\mathbf{1} + f) + (\beta\mathbf{1} + g) = (\alpha + \beta)\mathbf{1} + (f + g),$$

$$(\alpha\mathbf{1} + f) \cdot (\beta\mathbf{1} + g) = (\alpha\beta)\mathbf{1} + (\beta f + \alpha g + f \cdot g).$$

Then R is a commutative Banach algebra with unity. When G is discrete, $R = L_1(G)$. The Banach algebra R is called the **group algebra** of G . The group algebra R of G is † semisimple. R is algebraically isomorphic to a subalgebra of $C(\mathfrak{M})$, which is the associative algebra of all continuous functions on the compact Hausdorff space \mathfrak{M} consisting of all maximal ideals in R (\rightarrow 36 Banach Algebras). By this correspondence, if $\varphi \in R$ corresponds to $\varphi(M)$, which is a function on \mathfrak{M} , then $\sup_{M \in \mathfrak{M}} |\varphi(M)| \leq \|\varphi\|$. $L_1(G)$ belongs to \mathfrak{M} if and only if the group G is not discrete.

I. Fourier Transforms

According as the group G is discrete or not, we set $\mathfrak{R} = \mathfrak{M}$ or $\mathfrak{R} = \mathfrak{M} - \{L_1(G)\}$. Then there exists a one-to-one correspondence between the elements $M \in \mathfrak{R}$ and the elements χ of the † character group \hat{G} of G such that the following formulas are valid:

$$f(M) = \int_G \chi(x) f(x) dx, \quad f \in L_1, \quad (4)$$

$$\chi(y) = f_y(M) / f(M), \quad (5)$$

where $f_y(x) = f(xy^{-1})$ and f is a function such that $f(M) \neq 0$. This correspondence $M \leftrightarrow \chi$ gives a homeomorphism between the locally compact space \mathfrak{R} and \hat{G} . Hence if we identify M with χ and set $f(M) = \hat{f}(\chi)$, then \hat{f} is a continuous function on \hat{G} , called the **Fourier transform** of $f(x)$. Since $f \rightarrow f(M)$ is an algebraic isomorphism, $(fg)^{\wedge}(\chi) = \hat{f}(\chi)\hat{g}(\chi)$. If $\hat{f}(\chi) \equiv \hat{g}(\chi)$, then f is equal to g in $L_1(G)$. This is the **uniqueness theorem** of Fourier transforms. From it we can deduce the † maximal almost periodicity of locally compact Abelian groups.

If G is not discrete, then $L_1(G) \in \mathfrak{M}$, and $f(L_1(G)) = 0$ for $f \in L_1(G)$. Hence $\{\chi \mid |\hat{f}(\chi)| \geq \varepsilon\}$ is a compact subset of \hat{G} . This means that \hat{f} is a continuous function vanishing at infinity on \hat{G} . (This is a generalization of the † Riemann-Lebesgue theorem concerning the cases $G = \mathbf{R}^1$ or $\mathbf{T}^1 = \mathbf{R}/\mathbf{Z}$.) Any continuous function $u(\chi)$ on \hat{G} vanishing at infinity is approximated uniformly by $\hat{f}(\chi)$, which is the Fourier transform of $f \in L_1(G)$.

J. Positive Definite Functions

A function $\varphi(x)$ defined on G is said to be **positive definite** (or of **positive type**) if the inequality $\sum_{j,k=1}^n \varphi(x_j x_k^{-1}) \alpha_j \bar{\alpha}_k \geq 0$ holds for arbitrary elements x_1, \dots, x_n in G and arbitrary complex numbers $\alpha_1, \dots, \alpha_n$. We denote by P_G the set of all positive definite functions on G . If $\varphi \in P_G$, then $\varphi(e) > 0$ (e is the identity of G), $|\varphi(x)| \leq \varphi(e)$, and $\varphi(x^{-1}) = \overline{\varphi(x)}$. If G is locally compact, we further assume that $\varphi \in P_G$ is measurable with respect to the Haar measure of G . Then for any $f \in L_1(G)$,

$$\iint \varphi(xy^{-1}) f(y) \overline{f(y)} dx dy \geq 0.$$

Any $\varphi \in P_G$ is equal almost everywhere to a continuous $\varphi_1 \in P_G$. (Concerning positive definite functions on locally compact groups and their relation with unitary representations \rightarrow 437 Unitary Representations B.)

K. Harmonic Analysis and the Duality Theorem

If G is a locally compact Abelian group, then a function $\varphi(x)$ on G belongs to P_G if and only if there exists a nonnegative measure $\mu(\hat{G}) < \infty$ on \hat{G} such that $\varphi(x) = \int_G \chi(x) d\mu(\chi)$. (When $G = \mathbf{R}^1$, this theorem is † Bochner's theorem, whereas when $G = \mathbf{Z}$, it is Herglotz's theorem.) Hence we can prove a spectral resolution of unitary representations of G : $U(x) = \int_G \chi(x) dE(\chi)$ (a generalization of † Stone's theorem). If $f \in L_1(G) \cap P_G$, then $\hat{f}(\chi) \geq 0$, $\hat{f} \in L_1(\hat{G})$, and the **inversion formula** of Fourier transforms $f(x) = \int_G \bar{\chi}(x) \hat{f}(\chi) d\chi$ holds, provided that the Haar measure on G is suitably chosen. If $f \in L_1(G) \cap L_2(G)$, then $\hat{f} \in L_2(\hat{G})$, and **Parseval's identity** $\int_G |f(x)|^2 dx = \int_G |\hat{f}(\chi)|^2 d\chi$ holds. If we put $Uf = \hat{f}$ and $V\hat{f} = f$, then U is extendable uniquely to an isometry of $L_2(G)$ onto $L_2(\hat{G})$ and V is extendable uniquely to its inverse transformation, respectively (**Plancherel's theorem** on locally compact Abelian groups). By the inversion formula and Plancherel's theorem, we can prove that the character group $\hat{\hat{G}}$ of \hat{G} is isomorphic to G as a topological group. This is called the **Pontryagin duality theorem** of locally compact Abelian groups (\rightarrow 422 Topological Abelian Groups). In particular, if G is compact, \hat{G} is a discrete Abelian group. Then we can normalize the Haar measure of G and \hat{G} so that the measure of G and the measure of each element of \hat{G} are 1. Plancherel's theorem implies that the set of characters of G is a † complete † orthonormal set in $L_2(G)$.

L. Poisson's Summation Formula

Suppose that G is a locally compact Abelian group, H is its discrete subgroup, and G/H is compact. Then the * annihilator Γ of H is a discrete subgroup of \hat{G} . For any continuous function $f(x)$ on G , if $\sum_{y \in H} f(xy)$ is convergent absolutely and uniformly (hence $f \in L_1(G)$) and $\sum_{\xi \in \Gamma} \hat{f}(\xi)$ is convergent absolutely, then $\sum_{y \in H} f(y) = c \sum_{\xi \in \Gamma} \hat{f}(\xi)$, where c is a constant depending on the Haar measures of G and \hat{G} . This is called **Poisson's summation formula** on a locally compact Abelian group (a generalization of * Poisson's summation formula on $G = \mathbf{R}$).

M. Closed Ideals in $L_1(G)$

In the following, the group operations in G and \hat{G} are denoted by $+$, and the value of the character $\gamma(x)$ ($x \in G, \gamma \in \hat{G}$) is written as (x, γ) .

For a function f defined on G and any $y \in G$, the translation operator τ_y is defined by $\tau_y f(x) = f(x - y)$. A closed subspace of $L_1(G)$ is an ideal in $L_1(G)$ if and only if it is invariant under all translations (N. Wiener). A closed ideal I coincides with $L_1(G)$ if and only if the set of zeros of I , i.e., $Z(I) = \bigcap_{f \in I} \hat{f}^{-1}(0)$, is empty (Wiener's Tauberian theorem). A closed ideal I is maximal if and only if $Z(I)$ consists of a single point. If the dual \hat{G} of G is discrete, then the closed ideals in $L_1(G)$ are completely characterized by the zeros; that is, * spectral synthesis is possible, but this situation does not hold generally. P. Malliavin's theorem states that if \hat{G} is not discrete, then there exists a set E in \hat{G} and two different closed ideals I and J such that $Z(I) = Z(J) = E$. Such a set E is called a non-S-set. For example, if $\hat{G} = \mathbf{R}^3$, the unit sphere is a non-S-set (L. Schwartz).

N. Operating Functions

Denote by $A(\hat{G})$ the set of all Fourier transforms of the functions in $L_1(G)$. Let $\hat{f} \in A(\hat{G})$ and Φ an analytic function in a neighborhood of the range of \hat{f} . Furthermore, assume that $\Phi(0) = 0$ if G is not discrete. Then there exists a $\hat{g} \in A(\hat{G})$ such that $\hat{g}(\gamma) = \Phi(\hat{f}(\gamma))$ for $\gamma \in \hat{G}$ (Wiener-Lévy theorem). In general a function Φ defined on a set D in the complex plane is said to **operate** in a function algebra R or to be an **operating function** on R if $\Phi(f) \in R$ for all $f \in R$ whose range lies in D . The converse of the Wiener-Lévy theorem holds in the following form. Let G be an infinite Abelian group and Φ a function on the interval $[-1, 1]$. If Φ operates in $A(G)$, then Φ is analytic in a neighborhood of the origin if G is compact, and

analytic in a neighborhood of $[-1, 1]$ if G is not compact [14, 15]. Let E be a subset of \hat{G} , and I_E the set of all $\hat{f} \in A(\hat{G})$ such that $\hat{f} = 0$ on E . Let $A(E) = A(\hat{G})/I_E$ be the quotient algebra. The set E is called a **set of analyticity** if every operating function on $A(E)$ is analytic. For a characterization of such a set \rightarrow [11].

O. Measure Algebras

For a locally compact Abelian group G , let $M(G)$ be the set of all * regular bounded complex measures. For λ and μ in $M(G)$, the convolution $\lambda * \mu$ is defined by $(\lambda * \mu)(E) = \int_G \lambda(E - y) \cdot d\mu(y)$, where E is a Borel set in G . Then $M(G)$ is a semisimple commutative Banach algebra whose product is defined to be the convolution. The Fourier-Stieltjes transform of $\mu \in M(G)$ is defined by

$$\hat{\mu}(\gamma) = \int_G (x, \gamma) d\mu(x), \quad \gamma \in \hat{G}.$$

A continuous function on \hat{G} is positive definite if and only if it is the Fourier-Stieltjes transform of a positive measure in $M(G)$ (Bochner's theorem). Assume that G is not discrete. A function on the interval $[-1, 1]$ that operates in the Fourier-Stieltjes transforms of measures in $M(G)$ can be extended to an entire function, and a function on the whole complex plane that operates in the * Gelfand representation of $M(G)$ is an entire function [15]. From this fact, it follows that $M(G)$ is asymmetric and nonregular. Furthermore, there exists a measure $\mu \in M(G)$ such that $\hat{\mu}(\gamma) \geq 1$ but $1/\hat{\mu}$ is not a Fourier-Stieltjes transform of $M(G)$. (See also Wiener and Pitt [16], Shreider [17], and Hewitt and Kakutani [18]; for the general description of measure algebra, see Rudin [11] and Hewitt and Ross [12].)

P. Idempotent Measures

A measure $\mu \in M(G)$ is called **idempotent** if $\mu * \mu = \mu$, that is, $\hat{\mu}(\gamma) = 0$ or 1 for all $\gamma \in \hat{G}$. Then $\hat{\mu}$ is the characteristic function of the set $\{\gamma \in \hat{G} | \hat{\mu}(\gamma) = 1\}$. The smallest ring of subsets of \hat{G} that contains all open cosets of subgroups of \hat{G} is called the coset ring of \hat{G} . The characteristic function of a set E in \hat{G} is the Fourier-Stieltjes transform of an idempotent measure in $M(G)$ if and only if E belongs to the coset ring of \hat{G} [19]. A simple proof of this theorem is given by T. Ito and I. Amemiya (*Bull. Amer. Math. Soc.*, 70 (1964)). When G is the unit circle, the coset ring consists of sequences periodic except at a finite number of points, and for this case the theorem was obtained by H. Helson. Let

n_1, n_2, \dots, n_k be distinct integers and $d\mu(x) = \sum_{j=1}^k e^{in_j x} \cdot dx$. Then μ is an idempotent measure on the unit circle. J. E. Littlewood conjectured that the norm of μ exceeds $c \log k$, where c is a positive constant not depending on the choice of $\{n_j\}$. A partial answer was given by P. J. Cohen [19] for compact connected Abelian groups and was improved by H. Davenport (*Mathematika*, 7 (1960)) and E. Hewitt and H. S. Zuckerman (*Proc. Amer. Math. Soc.*, 14 (1963)).

Q. Mappings of Group Algebras

Let G and H be two locally compact Abelian groups and φ a nontrivial homomorphism of $L_1(G)$ into $M(H)$. Associated with φ there is a mapping φ^* of a subset Y of \hat{H} into \hat{G} such that $\varphi(f)(\gamma) = \hat{f}(\varphi^*(\gamma))$ for $\gamma \in Y$ and $= 0$ for $\gamma \notin Y$, or symbolically $\varphi(f) = \hat{f}(\varphi^*)$. A continuous mapping α of Y into \hat{G} is said to be piecewise affine if there exist a finite number of mutually disjoint sets S_j , $j = 1, \dots, n$, in the coset ring of \hat{H} and mappings α_j such that (i) $Y = \bigcup_{j=1}^n S_j$; (ii) α_j is defined on an open coset K_j of \hat{H} , where $K_j \supset S_j$; (iii) $\alpha_j = \alpha$ on S_j ; and (iv) $\alpha_j(\gamma + \gamma' - \gamma'') = \alpha_j(\gamma) + \alpha_j(\gamma') - \alpha_j(\gamma'')$ for all $\gamma, \gamma', \gamma'' \in K_j$, $j = 1, \dots, n$. P. J. Cohen's theorem is: If φ is a homomorphism of $L_1(G)$ into $M(H)$, then Y belongs to the coset ring of \hat{H} and φ^* is a piecewise affine mapping of Y into \hat{G} . Conversely, for any piecewise affine mapping α , there is a homomorphism φ of $L_1(G)$ into $M(H)$ such that $\varphi^* = \alpha$. Related theorems have been studied by A. Beurling, H. Helson, J.-P. Kahane, Z. L. Leibenson, and W. Rudin [11].

R. Exceptional Sets

Let G be a locally compact Abelian group. A subset E is said to be independent if $n_1 x_1 + \dots + n_k x_k = 0$, where the n_j are integers and $x_j \in E$ implies $n_j x_j = 0$, $j = 1, \dots, k$. A set E in G is called a **Kronecker set** if for every continuous function φ on E of absolute value 1 and $\varepsilon > 0$ there exists a $\gamma \in \hat{G}$ such that $|\varphi(x) - (x, \gamma)| < \varepsilon$, $x \in E$. Every Kronecker set is independent and of infinite order, but independent sets are not necessarily Kronecker sets. For a group G whose elements are of finite order p , a set E is called of type K_p if for every continuous function φ on E with values $\exp(2\pi i k/p)$, $k = 0, \dots, p-1$, there is a $\gamma \in \hat{G}$ such that $\varphi = \gamma$ on E . If E is a compact Kronecker set in G and μ is a measure with support in E , i.e., $\mu \in M(E)$, then $\|\mu\| = \|\hat{\mu}\|_\infty$. A compact set E is called a **Helson set** if there is a constant C such that $\|\mu\| \leq C \|\hat{\mu}\|_\infty$ for $\mu \in M(E)$. Every K_p set is also a Helson set. For a Helson set E , $C(E) = A(E)$. A discrete

analog of a Helson set is a Sidon set. A subset F of a discrete group \hat{G} is called a **Sidon set** if there is a constant C such that $\sum_{\gamma \in F} |a_\gamma| \leq C \sup_x |\sum_{\gamma \in F} a_\gamma(x, \gamma)|$ for every polynomial $\sum a_\gamma(x, \gamma)$. For example, a lacunary sequence $\{n_k\}$, $n_{k+1}/n_k > q > 1$, of integers is a Sidon set. These sets are deeply connected with harmonic analysis on groups, and measures concentrated on these sets have some unexpected pathological properties (\rightarrow e.g., [11, 13]).

S. Tensor Algebras and Group Algebras

Let X and Y be compact Hausdorff spaces, and denote by $V(X, Y)$ the projective tensor product $C(X) \hat{\otimes} C(Y)$ of continuous function spaces $C(X)$ and $C(Y)$. The norm of $\varphi(x, y) = \sum_{j=1}^\infty f_j(x)g_j(y)$ is defined by $\|\varphi\| = \inf \sum_{j=1}^\infty \|f_j\|_\infty \|g_j\|_\infty$, where the infimum is taken for all expressions of φ . If \hat{G} is an infinite compact group, then there exist two subsets K_1 and K_2 such that (i) K_1 and K_2 are homeomorphic to the Cantor ternary set; (ii) the expression $\gamma_1 + \gamma_2$ of an element of $E = K_1 + K_2$ is unique, where $\gamma_1 \in K_1$ and $\gamma_2 \in K_2$; (iii) $K_1 \cap K_2 \neq \emptyset$; and (iv) $K_1 \cup K_2$ is a Kronecker set or a set of type K_p for some p . Varopoulos' theorem states that the algebra $V(K_1, K_2)$ is isomorphic to $A(E)$ which denotes the algebra of restriction of functions in $A(\hat{G})$ on the set E . By this theorem, the problems of spectral synthesis and operating functions of group algebras are transformed into problems of tensor algebras. For a more precise discussion \rightarrow [20].

References

- [1] E. C. Titchmarsh, Introduction to the theory of Fourier integrals, Clarendon Press, 1937.
- [2] A. Zygmund, Trigonometric series I, Cambridge Univ. Press, second edition, 1959.
- [3] N. Wiener, The Fourier integrals and certain of its applications, Cambridge Univ. Press, 1933.
- [4] R. E. A. C. Paley and N. Wiener, Fourier transforms in the complex domain, Amer. Math. Soc. Colloq. Publ., 1934.
- [5] K. Yosida, Functional analysis, Springer, 1965, sixth edition, 1980.
- [6] M. A. Naimark, Normed rings, Noordhoff, 1959. (Original in Russian, 1956.)
- [7] L. H. Loomis, An introduction to abstract harmonic analysis, Van Nostrand, 1953.
- [8] H. Cartan and R. Godement, Théorie de la dualité et analyse harmonique dans les groupes abéliens localement compacts, Ann. Sci. Ecole Norm. Sup., (3) 64 (1947), 79–99.

- [9] R. Godement, Théorèmes taubériens et théorie spectrale, *Ann. Sci. Ecole Norm. Sup.*, (3) 64 (1947), 118–138.
- [10] A. Weil, L'intégration dans les groupes topologiques et ses applications, *Actualités Sci. Ind.*, Hermann, 1940, second edition, 1951.
- [11] W. Rudin, Fourier analysis on groups, Interscience, 1962.
- [12] E. Hewitt and K. A. Ross, Abstract harmonic analysis, Springer, I, 1963; II, 1970.
- [13] J.-P. Kahane and R. Salem, Ensembles parfaits et séries trigonométriques, *Actualités Sci. Ind.*, Hermann, 1963.
- [14] Y. Katznelson, Sur le calcul symbolique dans quelques algèbres de Banach, *Ann. Sci. Ecole Norm. Sup.*, (3) 76 (1959), 83–123.
- [15] H. Helson, J.-P. Kahane, Y. Katznelson, and W. Rudin, The functions which operate on Fourier transforms, *Acta Math.*, 102 (1959), 135–157.
- [16] N. Wiener and H. R. Pitt, On absolutely convergent Fourier-Stieltjes transforms, *Duke Math. J.*, 4 (1938), 420–436.
- [17] Yu. A. Shreider (Šreider), The structure of maximal ideals in rings of measures with convolution, *Amer. Math. Soc. Transl.*, 81 (1953). (Original in Russian, 1950.)
- [18] E. Hewitt and S. Kakutani, Some multiplicative linear functionals on $M(G)$, *Ann. Math.*, (2) 79 (1964), 489–505.
- [19] P. J. Cohen, On a conjecture of Littlewood and idempotent measures, *Amer. J. Math.*, 82 (1960), 191–212.
- [20] N. T. Varopoulos, Tensor algebras and harmonic analysis, *Acta Math.*, 119 (1967), 51–112.

193 (X.29)

Harmonic Functions and Subharmonic Functions

A. General Remarks

A real-valued function u of \dagger class C^2 defined in a domain D in the n -dimensional Euclidean space \mathbf{R}^n is called **harmonic** if it satisfies the Laplace equation

$$\Delta u(P) = \frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_n^2} = 0 \quad (P = (x_1, \dots, x_n))$$

in D . A harmonic function is, by definition, twice continuously differentiable, but turns out to be real analytic. It is not true, however, that the solutions of the Laplace equation are real analytic. For example, for the function $u(x, y) = \operatorname{Re} \exp(-z^{-4})$ ($z = x + iy \neq 0$), $u(0, 0) = 0$, u_{xx} and u_{yy} exist everywhere, and u satisfies the

Laplace equation, but is not continuous at the origin.

The fundamental properties of harmonic functions do not depend essentially on n .

A real-valued function u of class C^2 satisfying the inequality $\Delta u \geq 0$ is called **subharmonic**. For a more general definition of subharmonic functions and their properties — Sections P–U.

B. Invariance of Harmonicity

Harmonicity in \mathbf{R}^2 is invariant under any \dagger conformal transformation. Namely, when there exists a conformal bijection sending a domain D in the xy -plane onto a domain D' in the $\xi\eta$ -plane, every harmonic function $u(x, y)$ on D is transformed into a harmonic function of (ξ, η) on D' . In \mathbf{R}^n for $n \geq 3$, harmonicity is not generally preserved under conformal transformations. However, harmonicity is preserved in the following special case: Let D be a domain in \mathbf{R}^n ($n \geq 3$), and consider the **inversion** $f: D \rightarrow D'$ defined by $f(x_1, \dots, x_n) = (x'_1, \dots, x'_n) = (a^2 x_1 / r^2, \dots, a^2 x_n / r^2)$, $r = (x_1^2 + \dots + x_n^2)^{1/2}$. Let $u(x_1, \dots, x_n)$ be a harmonic function on D , and let $v(x'_1, \dots, x'_n)$ be the function on D' obtained by applying the **Kelvin transformation** to u . Namely, $v(x'_1, \dots, x'_n) = (a/r')^{n-2} u(a^2 x'_1 / r'^2, \dots, a^2 x'_n / r'^2)$, where $r'^2 = x_1'^2 + \dots + x_n'^2$. Then the function v is harmonic on D' . A function u that is harmonic outside a compact set is called **regular at the point at infinity** if any Kelvin transform of u is harmonic in a neighborhood of the origin, in which case $u(P) \rightarrow 0$ as $\overline{OP} \rightarrow \infty$. Now let $T: x_k = x_k(x'_1, \dots, x'_n)$, $1 \leq k \leq n$, be a one-to-one analytic transformation of a domain D' onto another domain D . If there exists a positive function $\varphi(x'_1, \dots, x'_n)$ in D' such that $\varphi(x'_1, \dots, x'_n) u(x_1(x'_1, \dots, x'_n), \dots, x_n(x'_1, \dots, x'_n))$ is harmonic for any harmonic function $u(x_1, \dots, x_n)$ in D , then T is conformal. A conformal transformation as it is known in differential geometry is either (i) a \dagger similarity transformation, (ii) an inversion with respect to a sphere or a plane, or (iii) a finite combination of transformations of types (i) and (ii).

C. Examples of Harmonic Functions

(1) If u is a polynomial in x_1, \dots, x_n and harmonic in \mathbf{R}^n , then the terms of degree k in u form a harmonic function for each $k \geq 0$. A harmonic homogeneous polynomial is said to be a **spherical harmonic**. (2) $\log r$ in \mathbf{R}^2 and r^{2-n} in \mathbf{R}^n ($n \geq 3$) are harmonic except at $r = 0$. (3) Every \dagger logarithmic potential in \mathbf{R}^2 and every \dagger Newtonian potential in \mathbf{R}^n ($n \geq 3$) is harmonic outside the \dagger support of the measure. Con-

Harmonic Functions and Subharmonic Functions

versely, any harmonic function defined on a domain D is represented in an arbitrary relatively compact domain D' in D as the sum of a logarithmic ($n=2$) or Newtonian ($n \geq 3$) potential of a measure on $\partial D'$ and the potential of a \dagger double layer. (4) Both the real part u and the imaginary part v of an analytic function of a complex variable are harmonic. We call v a **conjugate harmonic function** of u . If u is harmonic on a \dagger simply connected domain D , then the conjugate v of u is given by

$$v(x, y) = \int_{(a, b)}^{(x, y)} \left(-\frac{\partial u}{\partial y} dx + \frac{\partial u}{\partial x} dy \right) + \text{constant},$$

where (a, b) is a fixed point in D and the path of integration is contained in D . When D is a \dagger multiply connected domain, v may take many values in accordance with the \dagger homology classes of the paths of integration.

D. Green's Formulas

In the following one should substitute "curve" for the term "surface" when $n=2$. Let D be a bounded domain whose boundary S consists of a finite number of closed surfaces that are piecewise of class C^1 . Let u and v be harmonic in D , and suppose that all the first-order partial derivatives of u and v have finite limits at every boundary point. We call D the inside of S . Let n be a normal on S toward the outside of D . Then the relation

$$\int_S \left(u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) d\sigma = 0 \quad (1)$$

follows immediately from \dagger Gauss's formula, where $d\sigma$ is the \dagger surface element on S . In particular, when v is identically equal to 1, formula (1) gives

$$\int_S \frac{\partial u}{\partial n} d\sigma = 0. \quad (2)$$

Equations (1) and (2) are called **Green's and Gauss's formulas**, respectively. Conversely, u is harmonic in D if u is a function of class C^2 in D and at every point $p \in D$, there is a sequence $\{r_k\}$ decreasing to zero and such that $\int_{S(P, r_k)} (\partial u / \partial n) d\sigma = 0$ ($k = 1, 2, \dots$), where $S(P, r_k)$ is the spherical surface with center at P and radius r_k . Another sufficient condition for u to be harmonic is that u is of class C^1 and, at every point $P \in D$, there be an $r_p > 0$ such that $\int_{S(P, r)} (\partial u / \partial n) d\sigma = 0$ for every r , $0 < r < r_p$ (Koebe, Bocher).

E. Mean Value Theorems

We assume that u is a harmonic function, D the domain of definition of u , and S the bound-

ary of D . The mean value of u on the surface or the interior of any ball in D is equal to the value of u at the center of the ball. Namely,

$$u(P) = \frac{1}{\tau_n r^n} \int_{B(P, r)} u d\tau = \frac{1}{\sigma_n r^{n-1}} \int_{S(P, r)} u d\sigma,$$

where τ_n and σ_n are the volume and surface area of a unit ball in \mathbf{R}^n , respectively, $B(P, r)$ is the open ball with center at P and radius r , and $d\tau$ is the volume element. These relations are called **mean value theorems**. Conversely, if v is continuous in D and at every point $P \in D$, there is a sequence $\{r_k\}$ decreasing to zero and such that the mean value of v over $B(P, r_k)$ or $S(P, r_k)$ is equal to $v(P)$ for each k , then v is harmonic in D . This result is called **Koebe's theorem**. From the mean value theorems the **maximum principle** for harmonic functions follows: Any nonconstant u assumes neither maximum nor minimum in D . If both u and v are harmonic in D and have the same finite boundary value at every point on S , then $u \equiv v$ in D by the maximum principle. This is called the **uniqueness theorem**.

F. Boundary Value Problems

The **first boundary value problem** (or **Dirichlet problem**) is the problem of finding a harmonic function defined on D that assumes boundary values prescribed on S (\rightarrow 120 Dirichlet Problem). The **second boundary value problem** (or **Neumann problem**) is the problem of finding a harmonic function u whose normal derivative $\partial u / \partial n$ is equal to a function f prescribed on the piecewise smooth boundary S . The solution, if it exists, is uniquely determined up to an additive constant. In order for the solution to exist, f should satisfy the condition $\int_S f d\sigma = 0$. The **third boundary value problem** is the problem of finding a harmonic function u on D that satisfies $\partial u / \partial n = hu + f$ on S , where h and f are functions prescribed on S . All these problems can be reduced to certain \dagger Fredholm integral equations. There is also the boundary value problem of mixed type, in which the boundary values are prescribed in a part of S and the normal derivatives are prescribed on the rest.

G. The Poisson Integral

Let D be a bounded domain with smooth boundary S and u a function harmonic in D and continuous on $D \cup S$. Let $G(P, Q)$ be Green's function in D . Then (1) yields

$$u(P) = -\frac{1}{\sigma_n} \int_S u(Q) \frac{\partial G(P, Q)}{\partial n_Q} d\sigma(Q).$$

In particular, if $D = B(0, r)$, then

$$u(P) = \frac{r^2 - \overline{OP}^2}{\sigma_n r} \int_{S(0, r)} \frac{u(Q)}{PQ^n} d\sigma(Q).$$

Conversely, given an integrable function f on $S(0, r)$, we set

$$u(P) = \frac{r^2 - \overline{OP}^2}{\sigma_n r} \int_{S(0, r)} \frac{f(Q)}{PQ^n} d\sigma(Q).$$

Then $u(p)$ is harmonic in $B(0, r)$ and converges to $f(Q)$ as P tends to any point Q on $S(0, r)$ where f is continuous. We call u a **Poisson integral**. Sometimes it is possible to represent a harmonic function u in $B(0, r)$ in the following form, which is more general than the Poisson integral:

$$u(P) = \frac{r^2 - \overline{OP}^2}{\sigma_n r} \int_{S(0, r)} \frac{1}{PQ^n} d\alpha(Q), \quad (3)$$

where α is a signed \dagger Radon measure on $S(0, r)$. In order for u to admit such a representation, it is necessary and sufficient that $\int_{S(0, r)} |u| d\sigma$ be a bounded function of r' for $0 < r' < r$, or equivalently, that the \dagger subharmonic function $|u|$ have a \dagger harmonic majorant. Furthermore, if α is absolutely continuous, then the Poisson integral representation of u is possible, and vice versa. A necessary and sufficient condition for the function u to admit the Poisson integral representation is that there exist a positive convex function $\varphi(t)$ on $t > 0$ such that $\varphi(t)/t \rightarrow \infty$ as $t \rightarrow \infty$ and $\varphi(|u|)$ has a harmonic majorant.

When D is a general domain in which Green's function exists, every positive harmonic function $u(P)$ is represented uniquely as the integral $\int K(P, Q) d\mu(Q)$, where $K(P, Q)$ is a \dagger Martin kernel and μ is a Radon measure on the Martin boundary B whose support is contained in a certain essential part of B , each point of which is called a minimal point. A similar integral representation appears in the theory of Markov processes (\rightarrow 260 Markov Chains I). The representation $\int K(P, Q) d\mu(Q)$ is a generalization of (3). In terms of a function similar to $\varphi(t)$, we can give a necessary and sufficient condition for $u(P)$ to be represented in the form $\int K(P, Q) f(Q) dv(Q)$, which corresponds to the Poisson integral representation and in which v is determined by $1 = \int K(P, Q) dv(Q)$. This condition is equivalent to the condition that $u(P)$ be **quasibounded**, i.e., that there exist an increasing sequence of bounded harmonic functions that converges to u [7].

A positive harmonic function u is said to be **singular** if any nonnegative harmonic minorant of u vanishes identically. Every positive harmonic function can be expressed as the sum of a quasibounded harmonic function and a singular one.

H. Expansion

Let $P_0 = (x_1^0, \dots, x_n^0)$ be a point in D , and denote the distance from P_0 to S by r . Then a harmonic function u is expanded uniquely into a power series

$$\sum a_{k_1, \dots, k_n} (x_1 - x_1^0)^{k_1} \dots (x_n - x_n^0)^{k_n}, \quad k_1, \dots, k_n \geq 0,$$

in $B(P_0, (\sqrt{2} - 1)r)$. Thus u is (real) analytic in D . If u vanishes on an open set in D , then $u \equiv 0$ in D . If the power series is written as $\sum_k h_k$ with spherical harmonics h_k of degree $k = 1, 2, \dots$, then this series converges over all of $B(P_0, r)$.

I. Sequences of Harmonic Functions

In this section, $\{u_m\}$ is a sequence of harmonic functions in a bounded domain D . First, if each u_m is bounded and continuous on $D \cup S$ and $\{u_m\}$ converges uniformly on S , then $\{u_m\}$ converges uniformly in D , and the limiting function u is harmonic in D . Moreover, $\partial^{k_1 + \dots + k_n} u_m / \partial x_1^{k_1} \dots \partial x_n^{k_n}$ converges to $\partial^{k_1 + \dots + k_n} u / \partial x_1^{k_1} \dots \partial x_n^{k_n}$ uniformly on any compact subset of D (**Harnack's first theorem**). Second, if $u_1 \leq u_2 \leq \dots$ in D and there is a point of D at which $\{u_m\}$ is bounded, then $\{u_m\}$ converges uniformly on any compact subset of D (**Harnack's second theorem**). The following **Harnack's lemma** is useful: If u is positive harmonic in D , P_0 is a point of D , and K is a compact subset of D , then there exist positive constants c and c' , depending only on P_0 and K , such that $cu(P_0) \leq u(P) \leq c'u(P_0)$ on K .

Any family of (locally) uniformly bounded harmonic functions is \dagger normal. A family of positive harmonic functions that is bounded at a point is also normal by Harnack's lemma. If $\int_D |u_k - u_m|^p d\tau \rightarrow 0$ as $k, m \rightarrow \infty$ for $p \geq 1$, then \dagger Hölder's inequality implies that $\{u_m\}$ converges uniformly on any compact subset of D . It follows that if $\int_D |\text{grad}(u_k - u_m)|^p d\tau \rightarrow 0$ as $k, m \rightarrow \infty$ and $\{u_n\}$ converges at a point in D , then there exists a harmonic function u in D such that $\int_D |\text{grad}(u_m - u)|^p d\tau \rightarrow 0$ as $m \rightarrow \infty$ and u_m converges to u uniformly on any compact subset of D , where P_0 is any point in D . Finally, if $\int_D |u_n|^p d\tau$ ($p \geq 1$) are bounded, then $\{u_n\}$ forms a normal family.

J. Level Surfaces and Orthogonal Trajectories

The set $\{P | u(P) = \text{constant}\}$ is called a **level surface (niveau or equipotential surface)**. When a is given as the constant, the level surface is called the a -level surface. Assume that u is not a constant. A point where $\text{grad } u$ vanishes is called **critical**. The set of critical points consists of at most countably many \dagger real analytic

manifolds of dimension $\leq n-2$ ($n = \dim D$, and a manifold of dimension 0 is understood to be a point). Any compact subset of D intersects only a finite number of such manifolds; we express this fact by saying that the manifolds do not cluster in D . Each of these manifolds is contained in a certain level surface. The complement of the critical points with respect to any level surface consists of real analytic manifolds of dimension $n-1$ that do not cluster in D .

For each noncritical point there exists an analytic curve passing through it such that $\text{grad } u$ is parallel to the tangent to the curve at each point on the curve. A maximal curve with this property is called an **orthogonal trajectory** (or **line of force**). Along every orthogonal trajectory, u increases strictly in one direction and hence decreases in the other, so that none of the orthogonal trajectories is a closed curve. There is exactly one orthogonal trajectory passing through any noncritical point. Therefore no two orthogonal trajectories intersect, and no orthogonal trajectory terminates at a noncritical point. Moreover, the set of limit points of any orthogonal trajectory in each direction does not contain any noncritical point. When u is a Green's function $G(P, Q)$, every orthogonal trajectory is called a **Green line**, and a Green line that originates at the pole Q and along which u decreases to 0 is called **regular**. For any sufficiently large a , the a -level surface Σ_a is an analytic closed surface homeomorphic to a spherical surface. Let E be a family of orthogonal trajectories originating at the pole. If the intersection A of E and a closed level surface Σ_a is an $(n-1)$ -dimensional measurable set, then the harmonic measure of A at Q with respect to the interior of Σ_a is called the **Green measure** of E . M. Brelot and G. Choquet proved that all orthogonal trajectories originating at the pole except those belonging to a family of Green measure zero are regular. Consider a domain D bounded by two compact sets, and denote by u the harmonic measure of one compact set with respect to D . Assume that u is not a constant. Then u changes from 0 to 1 along all orthogonal trajectories except those belonging to a family that is small with respect to a measure similar to the Green measure (see "flux" defined in Section K).

K. Harmonic Flows

Denote by Σ_a the a -level surface for a harmonic function u , and by Σ_a^0 the complement of the set of critical points with respect to Σ_a . Let σ be an $(n-1)$ -dimensional domain in Σ_a^0 such that the $(n-2)$ -dimensional boundary of σ is piecewise of class C^2 . Suppose that u assumes

the value b ($> a$) on each orthogonal trajectory passing through σ . Consider the union of orthogonal trajectories that pass through σ . The subset of this union on which u assumes values between a and b forms a set called a **regular tube**. The parts of the boundary corresponding to a and b are called the lower and upper bases of the tube; accordingly, σ is the lower base. The integral $\int (\partial u / \partial n) d\sigma$ on any section (i.e., the part of a level surface in the tube) is constant and is called the **flux** of the tube. The family of orthogonal trajectories passing through an $(n-1)$ -dimensional domain (not necessarily bounded by a smooth boundary) in Σ_a^0 is called a **harmonic flow**, and a subfamily is called a harmonic subflow if its intersection A with Σ_a^0 is measurable (in the $(n-1)$ -dimensional sense). The flux of a harmonic subflow is defined to be $\int_A (\partial u / \partial n) d\sigma$. Then the Green measure of family E of Green lines originating at the pole is equal to the flux of E divided by σ_n . We can compute the exact value of the extremal length of any harmonic subflow.

L. Isolated Singularities

Let u be harmonic in an open ball except at the center O . It is expressed as the sum of a function $h(P)$ harmonic in the entire ball and $\sum_{m=0}^{\infty} H_m(P) / \overline{OP}^{2m+1}$, where H_m is a spherical harmonic of degree m . If $\overline{OP}^\alpha u(P) \rightarrow 0$ as $P \rightarrow O$ for $\alpha > 0$, then $u(P)$ is equal to $h(P) + c / \overline{OP} + \dots + H_m(P) / \overline{OP}^{2m+1}$ with $m < \alpha - 1$. In particular, if u is bounded in a neighborhood of O , then O is a **removable singularity** for u . If u is bounded above (below), then $u(P) = h(P) + c / \overline{OP}^{n-2}$, where $c \leq 0$ ($c \geq 0$). (When $n = 2$, we have $u(p) = h(p) + c \log 1 / \overline{OP}$. For the removability of a set of capacity zero \rightarrow 169 Function-Theoretic Null Sets).

If u is harmonic near the point at infinity, i.e., outside some closed ball, then

$$u(P) = \sum_{m=0}^{\infty} \frac{H_m(P)}{\overline{OP}^{2m+1}} + \sum_{m=0}^{\infty} U_m(P),$$

where the first sum is regular at the point at infinity and U_m is a spherical harmonic of degree m . If $\overline{OP}^{-\alpha} u(P) \rightarrow 0$ as $\overline{OP} \rightarrow \infty$ with $\alpha \geq 0$, then $U_m \equiv 0$ for all $m > \alpha$. If u is bounded above or below, then $U_m \equiv 0$ for all $m \geq 1$. If u is harmonic in \mathbf{R}^n and $\overline{OP}^{-\alpha} u(P) \rightarrow 0$ as $\overline{OP} \rightarrow \infty$ with $\alpha \geq 0$, then u is a polynomial of degree m ($< \alpha$). If u is harmonic and bounded above or below in \mathbf{R}^n , then u is constant. Brelot called a function u harmonic at the point at infinity if

$$u(P) = \text{constant} + \sum_{m=1}^{\infty} \frac{H_m(P)}{\overline{OP}^{2m+1}}$$

(note that $m \geq 1$) near the point at infinity [1].

M. Harmonic Continuation

If u vanishes in a subdomain of D , then $u \equiv 0$ in D . If u_k is harmonic in D_k ($k = 1, 2$), $D_1 \cap D_2 \neq \emptyset$, and $u_1 \equiv u_2$ in $D_1 \cap D_2$, then u_1 and u_2 define a harmonic function in $D_1 \cup D_2$. If the boundaries of mutually disjoint domains D_1 and D_2 have a surface S_0 of class C^1 in common, u_k is harmonic in D_k ($k = 1, 2$), $u_1 = u_2$ on S_0 , and $\partial u_1 / \partial n$ and $-\partial u_2 / \partial n$ exist and coincide on S_0 , then u_1 and u_2 define a harmonic function in the domain $D_1 \cup S_0 \cup D_2$. We express this fact by saying that one of u_1 and u_2 is a **harmonic continuation** of the other. It follows that $u \equiv 0$ in D if the boundary of D contains a surface S_0 of class C^1 and $u = \partial u / \partial n = 0$ on S_0 . Consider the case $n = 2$. If the boundary of a Jordan domain contains an analytic arc C and u (or $\partial u / \partial n$) vanishes on C , then a harmonic continuation of u into a certain domain beyond C is possible. If $n = 3$, however, nothing is known except in the case where S_0 is a part of a spherical surface or a plane and $u = 0$ (or $\partial u / \partial n = 0$) on S_0 .

Boundary values of u do not always exist, but in some special cases, u has limits. For instance, a positive harmonic function in a ball has a finite limit at almost every boundary point Q if the variable is restricted to any angular domain with vertex at Q .

N. Green Spaces

As a generalization of Riemann surfaces, Brelot and Choquet introduced \mathcal{E} -spaces [3]. It is required that \mathcal{E} be a separable connected topological space and satisfy the following two conditions: (i) At each point P there exists a neighborhood V_P of P and a homeomorphism between V_P and an open set V'_P in the Alexandrov compactification $\mathbf{R}^n \cup \{\infty\}$; (ii) if $A = V_{P_1} \cap V_{P_2} \neq \emptyset$ and A'_k is the part of V'_{P_k} that corresponds to A ($k = 1, 2$), then the correspondence between A'_1 and A'_2 via A is a conformal (possibly with the sense of angles reversed) transformation when $n = 2$ and an isometric transformation (which keeps ∞ invariant) when $n \geq 3$. If a Green's function exists on \mathcal{E} , then \mathcal{E} is called a **Green space**. Harmonic functions and the Dirichlet problem on a Green space have been discussed from various points of view.

O. Biharmonic Functions

A function v is called **polyharmonic** if $\Delta^k v = 0$ ($k \geq 2$) and **biharmonic** if $\Delta \Delta v = 0$; sometimes, polyharmonic functions are also called biharmonic. A biharmonic function in a plane domain D is written as $\operatorname{Re}(\bar{z}f(z) + g(z))$, where f

and g are complex analytic in D (Goursat's representation). Biharmonic functions are used in the theory of elasticity and hydrodynamics.

P. Subharmonic Functions

Let D be a domain in the n -dimensional Euclidean space \mathbf{R}^n ($n \geq 2$). A real-valued function $u(P)$ in D is called **subharmonic** if (1) $-\infty \leq u < +\infty$, $u \neq -\infty$; (2) u is upper semicontinuous; and (3) at every point P_0 of D , the mean value of u over the surface of any closed ball in D with center at P_0 is not smaller than $u(P_0)$, i.e.,

$$u(P_0) \leq \frac{1}{\sigma_n r^{n-1}} \int u d\sigma \equiv L(P_0, r),$$

where σ_n is the area of the surface of a unit ball in \mathbf{R}^n . Condition (3) can be replaced by: (3') The mean value $A(P_0, r)$ of u over the closed ball is $\geq u(P_0)$. In order that an upper semicontinuous function u be subharmonic it is necessary and sufficient that, for any subdomain D' of D and for any harmonic function h in D' , the maximum principle hold for $u - h$.

We call $-u$ **superharmonic** when u is subharmonic. A harmonic function is subharmonic and superharmonic. The converse is also true (\rightarrow Section E).

When u is of class C^2 , then u is subharmonic if and only if

$$\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_n^2} \geq 0 \quad (P = (x_1, \dots, x_n)).$$

When u is an upper semicontinuous function that is not necessarily differentiable, u is subharmonic if and only if Δu interpreted as a distribution is a positive measure.

If u_1, \dots, u_k are subharmonic and a_1, \dots, a_k are positive constants, then $a_1 u_1 + \dots + a_k u_k$ and $\max(u_1(P), u_2(P), \dots, u_k(P))$ are subharmonic. If a subharmonic function u is replaced by the Poisson integral for the boundary function u inside a closed ball in D , then the resulting function in D is subharmonic. If $f(t)$ is a monotone increasing convex function, then $f(u)$ is subharmonic. If $v \geq 0$ and $\log v$ is subharmonic, then v is subharmonic. If $f(z)$ is a holomorphic function of the complex variable z and $\lambda > 0$, then $\lambda \log |f(z)|$ and hence $|f(z)|^\lambda$ are subharmonic. If h is harmonic, then $|h|$ is subharmonic. Any logarithmic potential ($n = 2$) or Newtonian potential ($n \geq 3$) is superharmonic in \mathbf{R}^n .

Q. Properties of Mean Values

Condition (3) (resp. (3')) can be replaced by the condition that there exists an $r(P_0) > 0$ at any

Harmonic Functions and Subharmonic Functions

P_0 such that $u(P_0) \leq L(P_0, r)$ ($u(P_0) \leq A(P_0, r)$) for every r , $0 < r < r(P_0)$. The relation $-\infty < A(P_0, r) \leq L(P_0, r)$ always holds, and both $A(P_0, r)$ and $L(P_0, r)$ decrease to $u(P_0)$ as $r \downarrow 0$. On any \dagger compact subset of D , u is \dagger integrable. Both $A(P_0, r)$ and $L(P_0, r)$ increase with r and are convex functions of $-\log r$ ($n=2$) and r^{2-n} ($n \geq 3$); hence they are continuous functions of r . If D' is relatively compact in D and r_0 is the distance between $\partial D'$ and ∂D , then $A(P, r)$ is a continuous subharmonic function of P in D' , where r is fixed in the interval $(0, r_0)$. By taking the average of $A(P, r)$ k times, a subharmonic function of class C^k is obtained that decreases to u as $r \downarrow 0$. If $\varphi_r((x_1^2 + \dots + x_n^2)^{1/2})$ is suitably chosen, then the \dagger convolution $u * \varphi_r$ is a subharmonic function of class C^∞ and decreases to u as $r \downarrow 0$.

R. Sequences of Subharmonic Functions

The limit of a decreasing sequence or a downward-directed net of subharmonic functions is subharmonic or equal constantly to $-\infty$. The limit of a uniformly convergent sequence of subharmonic functions is subharmonic. If u_1, u_2, \dots are subharmonic, then $\max(u_1, \dots, u_k)$ is subharmonic for every k , but $\sup(u_1, u_2, \dots)$ may not be subharmonic. Let U be a family of subharmonic functions in D that are uniformly bounded above on every compact subset of D . Then the **upper envelope** of U , i.e., the function defined by $\sup_{u \in U} u$ in D , coincides with a subharmonic function except on a set of \dagger capacity zero.

S. Harmonic Majorants and Riesz Decompositions

Suppose that we are given a subharmonic function u in D . If there is a harmonic function h satisfying $h \geq u$ in D , then h is called a **harmonic majorant** of u . When there is a harmonic majorant of u , there exists a least one among them, denoted by h_D . For any relatively compact subdomain D' of D , $h_{D'}$ always exists and equals the \dagger Perron-Brelot solution in D' for the boundary function u . As D' increases to D , $h_{D'}$ increases to a function h that is either harmonic or equal constantly to ∞ . If h_D exists, it coincides with h , and hence h is harmonic. Conversely, if h is harmonic, then h_D exists and equals h . Generally, there is a unique \dagger Radon measure μ in D with the following property: Let δ be any subdomain of D such that h_δ and the \dagger Green's function G_δ exist in δ (δ may coincide with D). Then $h_\delta - u$ is equal to the potential $\int_\delta G_\delta d\mu$, and $u = h_\delta - \int_\delta G_\delta d\mu$. In general, a representation of a superharmonic (subharmonic) function as the sum (difference)

of a harmonic function and a potential is called a **Riesz decomposition**.

T. Boundary Values

Let D be a domain in which a Green's function exists, and consider the \dagger fine topology on the \dagger Martin compactification of D . Any negative subharmonic function u in D has a finite limit with respect to the fine topology at every point of the \dagger Martin boundary Δ except at the points of a subset of Δ of harmonic measure zero (J. L. Doob). When D is a ball, u has a limit in the ordinary sense along almost every radius. However, it may happen that even if u is bounded there exists no angular limit at any point of the boundary. If the mean value of $|u|$ on every concentric smaller ball in D is bounded, then u can be decomposed into the sum of a nonnegative harmonic function and a negative subharmonic function by the Riesz decomposition, and hence u has a limit both radially and with respect to the fine topology.

Let D be a domain and K be a compact subset of D of \dagger capacity zero. If a function u is subharmonic and bounded above in $D - K$, then u can be extended to be a subharmonic function in D . A function that is equal to a subharmonic function almost everywhere is called **almost subharmonic**, and an almost subharmonic function satisfying condition (3') is called **submedian**.

Subharmonic functions can be discussed in a space more general than \mathbf{R}^n , e.g., a \dagger Riemann surface ($n=2$), or more generally, an $\dagger\mathcal{E}$ -space of dimension n (≥ 2) in the sense of Brelot and Choquet.

U. The Axiomatic Treatment

\dagger Newtonian potentials were the main object of interest in the early stages of \dagger potential theory. A major part of potential theory can be discussed on the basis of the theory of superharmonic functions [8]. For example, a \dagger polar set is defined as a set on which some superharmonic function assumes the value ∞ , and a set X is \dagger thin at a point $P_0 \notin X$ if and only if P_0 has a positive distance from X or there exists a superharmonic function $v(P)$ in a neighborhood of P_0 such that $\limsup v(P) > v(P_0)$ as $P \in X$ tends to P_0 . Moreover, we can discuss \dagger balayage, define potentials, and obtain Riesz decompositions. Generalizing results of Doob (1954) and starting from a family of harmonic functions defined axiomatically in a locally compact Hausdorff space, M. Brelot defined superharmonic functions and potentials and discussed balayage, Riesz decompositions, and

the \dagger Dirichlet problem (1957). Further progress in axiomatic potential theory has been made by Brelot, H. Bauer, C. Constantinescu, A. Cornea, and others [11, 14].

References

- [1] M. Brelot, Sur le rôle du point à l'infini dans la théorie des fonctions harmoniques, *Ann. Sci. Ecole Norm. Sup.*, (3) 61 (1944), 301–332.
- [2] M. Brelot, *Eléments de la théorie classique du potentiel*, Centre de Documentation Universitaire, Paris, third edition, 1965.
- [3] M. Brelot and G. Choquet, Espaces et lignes de Green, *Ann. Inst. Fourier*, 3 (1951), 199–263.
- [4] O. D. Kellogg, *Foundations of potential theory*, Springer, 1929.
- [5] M. Nicolescu, Les fonctions polyharmoniques, *Actualités Sci. Ind.*, Hermann, 1936.
- [6] M. Ohtsuka, Extremal length of level surfaces and orthogonal trajectories, *J. Sci. Hiroshima Univ.*, 28 (1964), 259–270.
- [7] M. Parreau, Sur les moyennes des fonctions harmoniques et analytiques et la classification des surfaces de Riemann, *Ann. Inst. Fourier*, 3 (1952), 103–197.
- [8] M. Brelot, Sur la théorie autonome des fonctions sous-harmoniques, *Bull. Sci. Math.*, (2) 65 (1941), 72–98.
- [9] M. Brelot, *Lectures on potential theory*, Tata Inst., 1960.
- [10] T. Radó, *Subharmonic functions*, *Erg. Math.*, Springer, 1937 (Chelsea, 1970).
- [11] Colloque International sur la Théorie du Potentiel, Paris-Orsay, 1964, C.N.R.S. (*Ann. Inst. Fourier*, 15 (1965), fasc. 1.)
- [12] L. Helms, *Introduction to potential theory*, Wiley Interscience, 1969.
- [13] W. K. Hayman and P. B. Kennedy, *Subharmonic functions I*, Academic Press, 1976.
- [14] C. Constantinescu and A. Cornea, *Potential theory on harmonic spaces*, Springer, 1972. Also → references to 120 Dirichlet Problem and 338 Potential Theory.

194 (VII.11) Harmonic Integrals

A. Introduction

\dagger De Rham's theorem shows that the cohomology group with real coefficients of a \dagger differentiable manifold of class C^∞ is isomorphic to the cohomology group of the cochain com-

plex of \dagger differential forms with respect to the exterior derivative d . Thus every element of the cohomology group can be represented by a class of \dagger closed differential forms. Harmonic forms enable us to choose one definite differential form in each cohomology class. The theory of harmonic forms, called the theory of **harmonic integrals**, is modeled after the theory of holomorphic differentials and their integrals (Abelian integrals) in function theory [2, 4, 5, 8].

B. Definitions

Let X be an oriented n -dimensional differentiable manifold of class C^∞ with a \dagger Riemannian metric ds^2 of class C^∞ (→ 105 Differentiable Manifolds, 364 Riemannian Manifolds A). For every (C^∞) p -form φ on X we define an $(n-p)$ -form $*\varphi$ on X as follows: First denote the \dagger volume element of X by dv . If we choose a basis $\{\omega_1, \dots, \omega_n\}$ of the space of 1-forms on an open set U of X such that $ds^2 = \sum_i \omega_i^2$ and $dv = \omega_1 \wedge \dots \wedge \omega_n$, then φ can be expressed on U in the form $\varphi = (1/p!) \sum \varphi_{i_1, \dots, i_p} \omega_{i_1} \wedge \dots \wedge \omega_{i_p}$. If we let $*\varphi = (1/(n-p)!) \sum (*\varphi)_{j_1, \dots, j_{n-p}} \omega_{j_1} \wedge \dots \wedge \omega_{j_{n-p}}$, where $(*\varphi)_{j_1, \dots, j_{n-p}} = (1/p!) \sum \delta_{i_1, \dots, i_p}^{j_1, \dots, j_{n-p}} \varphi_{i_1, \dots, i_p}$, then $*\varphi$ is an $(n-p)$ -form on U that does not depend on the choice of $(\omega_1, \dots, \omega_n)$ and is determined only by φ . Since X is covered by open sets as above, $*$ defines a linear mapping that transforms p -forms to $(n-p)$ -forms. If we let $ds^2 = \sum g_{jk} dx^j dx^k$ in terms of the local coordinate system (x^1, \dots, x^n) and $\varphi = (1/p!) \varphi_{i_1, \dots, i_p} dx^{i_1} \wedge \dots \wedge dx^{i_p}$, then, in the notation of tensor calculus, we have

$$*\varphi = (1/(n-p)!)(*\varphi)_{j_1, \dots, j_{n-p}} dx^{j_1} \wedge \dots \wedge dx^{j_{n-p}},$$

$$(*\varphi)_{j_1, \dots, j_{n-p}} = \sqrt{g} \cdot \delta_{k_1, \dots, k_p}^{j_1, \dots, j_{n-p}} \varphi^{k_1, \dots, k_p}$$

$$(g = \det(g_{ij})).$$

For two p -forms φ and ψ , we define the inner product by $(\varphi, \psi) = \int_X \varphi \wedge *\psi$ if the right-hand side converges. In order for the inner product (φ, ψ) to be defined, it suffices that either φ or ψ has compact \dagger support. Then (φ, ψ) is a symmetric, positive definite bilinear form.

If we let $\delta = (-1)^{np+n+1} *d*$ operate on p -forms, where d is the \dagger exterior derivative, then d and δ are **adjoint** to each other with respect to the inner product. That is, if either φ or ψ has compact support, we have $(d\varphi, \psi) = (\varphi, \delta\psi)$ (\dagger Stokes's theorem). We call $\Delta = d\delta + \delta d$ the **Laplace-Beltrami operator**, which is a self-adjoint \dagger elliptic differential operator. These operators satisfy relations such as $** = (-1)^{p(n-p)}$, $dd = 0$, $\delta\delta = 0$, $*\Delta = \Delta*$,

$\ast \delta = (-1)^p d \ast$, and $\delta \ast = (-1)^{n-p+1} \ast d$ (when they operate on p -forms).

A differential form φ is said to be **harmonic** if $d\varphi = 0$ and $\delta\varphi = 0$. Then $\Delta\varphi = 0$. Since Δ is an elliptic operator, a \dagger weak solution φ of the equation $\Delta\varphi = \mu$ is an ordinary solution of class C^∞ on the domain where μ is of class C^∞ . Therefore, if φ is harmonic (as a weak solution), φ is of class C^∞ .

C. Harmonic Forms on Compact Manifolds

On a compact manifold X , any φ with $\Delta\varphi = 0$ is harmonic, since $(\varphi, \Delta\varphi) = (d\varphi, d\varphi) + (\delta\varphi, \delta\varphi)$. Let $L_p(X)$ be the linear space of p -forms of class C^∞ on X , and denote by $\mathfrak{L}_p(X)$ the completion of $L_p(X)$ with respect to the inner product (φ, ψ) . Then $\mathfrak{L}_p(X)$ is the Hilbert space of square integrable measurable p -forms. Then $\mathfrak{H}_p(X) = \{\varphi \in \mathfrak{L}_p(X) | \Delta\varphi = 0 \text{ (in the weak sense)}\}$ is a finite-dimensional subspace of $\mathfrak{L}_p(X)$ and is contained in $L_p(X)$, as we have seen before. Also, $\mathfrak{H}_p(X)$ is closed in $\mathfrak{L}_p(X)$, and the \dagger projection operator $H: \mathfrak{L}_p(X) \rightarrow \mathfrak{H}_p(X)$ is an \dagger integral operator with kernel of class C^∞ . The orthogonal complement of $\mathfrak{H}_p(X)$ in $\mathfrak{L}_p(X)$ is mapped onto itself by Δ and has the inverse operator G of Δ , which is a continuous operator of the Hilbert space. By letting $G = 0$ on $\mathfrak{H}_p(X)$, we can extend G to an operator from $\mathfrak{L}_p(X)$ to $\mathfrak{L}_p(X)$ that is called **Green's operator**. It is also denoted by G , maps L_p into itself, commutes with d and δ , and satisfies $GH = HG = 0$, $H + \Delta G = 1$ (= identity mapping). Therefore, for $\varphi \in L_p(X)$ we have $\varphi = H\varphi + G\delta d\varphi + d\delta G\varphi$, which shows that H is \dagger homotopic to the identity mapping of the \dagger cochain complex $(\sum_p L_p(X), d)$. From this we infer that every cohomology class of de Rham cohomology contains a unique harmonic form that represents the cohomology class. However, since products of harmonic forms are not always harmonic, it is not appropriate to use harmonic forms to study the ring structure of cohomology. G is also an integral operator with kernel of class C^∞ outside the diagonal subset in $X \times X$.

D. Harmonic Forms on Noncompact Manifolds

If X is a noncompact manifold, let $L_p(X)$ be the space of p -forms of class C^∞ with compact support, and let $\mathfrak{L}_p(X)$ be its completion. Let $\mathfrak{B}_p(X)$ and $\mathfrak{B}_p^\ast(X)$ be the respective closures of $dL_{p-1}(X)$ and $\delta L_{p+1}(X)$ in $\mathfrak{L}_p(X)$, and let $\mathfrak{Z}_p(X)$ and $\mathfrak{Z}_p^\ast(X)$ be the respective orthogonal complements of $\mathfrak{B}_p(X)$ and $\mathfrak{B}_p^\ast(X)$ in $\mathfrak{L}_p(X)$. Then $\mathfrak{Z}_p(X) \cap \mathfrak{Z}_p^\ast(X) = \mathfrak{H}_p(X)$ is a subspace of the square integrable harmonic forms, and we

have the direct sum decomposition $\mathfrak{L}_p(X) = \mathfrak{B}_p(X) + \mathfrak{B}_p^\ast(X) + \mathfrak{H}_p(X)$. In this decomposition any component of a form of class C^∞ is also of class C^∞ .

If X is an open submanifold of another manifold Y , \bar{X} is compact, and $\partial X = \bar{X} - X$ is a closed submanifold of Y , then the theory in this section is just a generalized potential theory with boundary condition $\varphi = 0$ on ∂X . We sometimes treat decompositions of other Hilbert spaces that correspond to other boundary conditions.

E. Generalization to Complex Manifolds

If X is a complex manifold, we consider complex-valued differential forms (\rightarrow 72 Complex Manifolds C). Then the space $L_p(X)$ of p -forms is the direct sum of the spaces $L_{r,s}(X)$ of forms of type (r, s) , and the exterior derivative d has the expression $d = d' + d''$, where d' is of type $(1, 0)$ (i.e., $L_{r,s}(X) \rightarrow L_{r+1,s}(X)$) and d'' is of type $(0, 1)$. If we are given a holomorphic vector bundle E on X , we can define an operator d'' on differential forms with values in E , and we have the generalized \dagger Dolbeault theorem. If X is compact and has a \dagger Hermitian metric, we can define a Hermitian inner product on E as follows: There is an open covering $\{U_j\}$ such that over each U_j the vector bundle E is isomorphic to $U_j \times C^q$. A point of E over U_j is represented by (x, ξ_j) , where $x \in U_j$ and $\xi_j \in C^q$. For $x \in U_j \cap U_k$ we have $(x, \xi_j) = (x, \xi_k)$ (the sides are the respective expressions over U_j and U_k) if and only if $\xi_j = g_{jk}(x)\xi_k$, where $g_{jk}(x)$ is a holomorphic mapping from $U_j \cap U_k$ to $GL(q, C)$ satisfying $g_{jk}g_{kl} = g_{jl}$ on $U_j \cap U_k \cap U_l$. A differential form φ with values in E is expressed as a family $\{\varphi_j\}$ of differential forms on U_j with values in C^q such that $\varphi_j(x) = g_{jk}(x)\varphi_k(x)$ on $U_j \cap U_k$. If we take a positive definite Hermitian matrix h_j whose components are C^∞ -functions on U_j such that ${}^t g_{jk} h_j \bar{g}_{jk} = h_k$ on $U_j \cap U_k$, then $\{\xi_j h_j \bar{\xi}_j\}$ determines a Hermitian inner product on each fiber of E . We can also endow the space $L_{r,s}(E, X)$ (of forms of type (r, s) of class C^∞ with values in E) with a Hermitian inner product by setting $(\varphi, \psi) = \int_X \sum_{\alpha, \beta} h_{j\alpha\beta} \varphi_j^\alpha \wedge \bar{\psi}_j^\beta$ for $\varphi, \psi \in L_{r,s}(E, X)$ (where the φ_j^α ($\alpha = 1, \dots, q$) are the components of φ_j). If we denote by \bar{d} the adjoint operator of d'' with respect to this inner product and let $A = d''\bar{d} + \bar{d}d''$, then A is a self-adjoint elliptic differential operator, and results similar to those for Δ mentioned above hold for A . For example, the space $\mathfrak{H}_{r,s}(E, X)$ of harmonic forms of type (r, s) is of finite dimension, and there is a continuous linear operator G on $\mathfrak{L}_{r,s}(E, X)$, the completion of $L_{r,s}(E, X)$, that satisfies $1 = H + AG$, $HG = GH = 0$, $d''G = Gd''$, and $\bar{d}G = G\bar{d}$. Here H

denotes the projection $\mathfrak{L} \rightarrow \mathfrak{H}$, which is an integral operator with kernel of class C^∞ . Also, G maps $L_{r,s}(E, X)$ into itself. Therefore H is homotopic to the identity on the cochain complex $(\sum_s L_{r,s}(E, X), d'')$, and any element of ${}^{\dagger}\text{Dolbeault's cohomology groups } (d''\text{-cohomology groups})$ is represented by a unique harmonic form (\rightarrow 232 Kähler Manifolds B).

F. Other Generalizations

Even if a manifold X is not of class C^∞ , if X is a manifold of class C^1_1 , we can develop the theory of harmonic forms [6]. We say that X is of class C^1_1 if it is of class C^1 and has a set of local coordinate systems whose transition functions have derivatives satisfying the ${}^{\dagger}\text{Lip-schitz condition}$.

If X is a real analytic manifold with a real analytic Riemannian metric, then harmonic forms are also real analytic. Using this fact, we can embed real analytically a compact manifold with a real analytic Riemannian metric into a Euclidean space (P. Bidal and G. de Rham; this result is now included in the theorems of C. B. Morrey and H. Grauert).

We can consider the theory of harmonic forms with singularities [4, 5], a generalization of the theory of differential forms of the second and third kinds. Here the notion of ${}^{\dagger}\text{current}$ is very useful.

G. Cohomology Vanishing Theorems

Since the operator Δ is closely related to the Riemannian metric, some metrics may admit no harmonic forms of certain degrees except zero. This is important since it means that the corresponding cohomology group of the manifold vanishes. The condition for this phenomenon to occur can be described in terms of the curvature of the metric. This study has its origin in S. Bochner's results [2].

Here is an example of a **vanishing theorem**: Let B be a holomorphic line bundle on a compact complex manifold X of dimension n . If the ${}^{\dagger}\text{Chern class of } B$ is expressed by a real closed differential form of type $(1, 1)$ as $\omega = \sqrt{-1} \sum a_{\alpha\beta} dz^\alpha \wedge d\bar{z}^\beta$, where the Hermitian matrix $(a_{\alpha\beta})$ is positive definite at every point of X , then $H^q(X, \Omega^p(B)) = 0$ for $p + q > n$. In this case, $ds^2 = 2 \sum a_{\alpha\beta} dz^\alpha \wedge d\bar{z}^\beta$ is a ${}^{\dagger}\text{Hodge metric on } X$ (\rightarrow 232 Kähler Manifolds D).

References

[1] W. L. Baily, The decomposition theorem for V -manifolds, Amer. J. Math., 78 (1956), 862–888.

195 B
Harmonic Mappings

[2] S. Bochner and K. Yano, Curvature and Betti numbers, Ann. Math. Studies 22, Princeton Univ. Press, 1953.
[3] S. I. Goldberg, Curvature and homology, Academic Press, 1962.
[4] W. V. D. Hodge, The theory and applications of harmonic integrals, Cambridge Univ. Press, second edition, 1952.
[5] K. Kodaira, Harmonic fields in Riemannian manifolds (generalized potential theory), Ann. Math., (2) 50 (1949), 587–665.
[6] C. B. Morrey and J. Eells, A variational method in the theory of harmonic integrals I, Ann. Math., (2) 63 (1956), 91–128.
[7] G. de Rham, Variétés différentiables, Actualités Sci. Ind., Hermann, second edition, 1960.
[8] G. de Rham and K. Kodaira, Harmonic integrals, Lecture notes, Inst. for Advanced Study, Princeton, 1950.

195 (VII.15)
Harmonic Mappings

A. General Remarks

The theory of harmonic mappings between Riemannian manifolds has its origin in the study of ${}^{\dagger}\text{Plateau's problem}$. The basic problem in the theory is to deform a given mapping into a harmonic one, which is a problem of the ${}^{\dagger}\text{calculus of variations and } {}^{\dagger}\text{global analysis}$ (\rightarrow 46 Calculus of Variations, 183 Global Analysis). Recently, the theory of harmonic mappings has been applied to problems in various branches of geometry [7–9, 11].

B. Definitions and Examples

Let (M, g) and (N, h) be ${}^{\dagger}\text{Riemannian manifolds with metrics } g = \sum g_{ij} dx^i dx^j$ and $h = \sum h_{\alpha\beta} dy^\alpha dy^\beta$, respectively. We define the **energy** of a C^1 -mapping $f: M \rightarrow N$ by

$$E(f) = \frac{1}{2} \int_M |df(x)|^2 dx,$$

where $|df(x)|$ is the ${}^{\dagger}\text{Hilbert-Schmidt norm of the differential } df_x: T_x(M) \rightarrow T_{f(x)}(N)$ of f at $x \in M$ and dx is the canonical Lebesgue measure defined by g on M (assumed compact). Thus $E(f)$ can be considered to be a generalization of the classical ${}^{\dagger}\text{Dirichlet integral for functions}$. The integrand $e(f)(x) = |df(x)|^2$ is called the **energy density** of f ; it measures the sum of the squares of elements of length stretched on a complete set of mutually perpendicular directions.

The \dagger Euler-Lagrange differential equations of the energy functional $E(f)$ yield a vector field $\tau(f)$ along f , i.e., a section of the bundle $f^*T(N)$ induced from the tangent bundle $T(N)$ of N by f . In fact, given a family f_t of mappings depending differentiably on t with $f_0 = f$, we have

$$\frac{d}{dt}E(f_t)|_{t=0} = - \int_M \left\langle \tau(f), \frac{\partial f_t}{\partial t} \Big|_{t=0} \right\rangle dx,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of tangent vectors along f . The vector field $\tau(f)$ is called the **tension field** of the mapping f ; it indicates the direction in which the energy of f decreases most rapidly.

The Euler-Lagrange differential equations $\tau(f) = 0$ are a system of \dagger quasilinear elliptic partial differential equations of the second order. In local coordinates, these can be written in the form

$$\Delta f^\alpha + \sum g^{ij} \Gamma_{\beta\gamma}^\alpha(f) \frac{\partial f^\beta}{\partial x^i} \frac{\partial f^\gamma}{\partial x^j} = 0,$$

where Δ is the \dagger Laplace-Beltrami operator on M and the $\Gamma_{\beta\gamma}^\alpha(f)(x)$ are the \dagger Christoffel symbols on N at $f(x)$. (The f^α are local coordinates of the point $f(x)$, and (g^{ij}) is the inverse matrix of (g_{ij}) .) A C^2 -mapping $f: M \rightarrow N$ is said to be **harmonic** if its tension field $\tau(f)$ vanishes. Thus, if M is compact, a C^2 -mapping $f: M \rightarrow N$ is harmonic if and only if it is an extremal of the energy functional $E(f)$.

Examples of harmonic mappings appear in various contexts of differential geometry. For instance:

- (1) If $N = \mathbf{R}$, then the harmonic mappings $M \rightarrow \mathbf{R}$ are the \dagger harmonic functions on M .
- (2) If M is the circle S^1 , then a harmonic mapping $S^1 \rightarrow N$ is a closed \dagger geodesic of N parametrized by arc length.
- (3) Let $f: M \rightarrow N$ be an \dagger isometric immersion of M into N . Then f is harmonic if and only if it is a \dagger minimal immersion.
- (4) If M and N are \dagger Kähler manifolds, then every \dagger holomorphic or **antiholomorphic** mapping $M \rightarrow N$ is harmonic, where by an antiholomorphic mapping is meant a mapping whose differential mapping carries a differential form of type $(1, 0)$ into that of type $(0, 1)$. We note that each (anti-)holomorphic mapping is an absolute minimum for the energy in its homotopy class. There are also examples of nonholomorphic (and nonantiholomorphic) harmonic mappings between Kähler manifolds (\rightarrow Section D).

C. Fundamental Properties

- (1) **Regularity.** Since a harmonic mapping is a solution of a second-order quasilinear elliptic

system of partial differential equations $\tau(f) = 0$ (\rightarrow Section B), it is a smooth (i.e., of class C^∞) mapping. More generally, it is known that a continuous mapping that satisfies $\tau(f) = 0$ in a weak sense is smooth [4, 6].

(2) **Unique continuation property.** The following unique continuation theorem is valid for harmonic mappings: If two harmonic mappings of M into N agree up to infinitely high order at some point of M , then they are identical (M being assumed connected). In particular, a harmonic mapping that is constant on an open set is a constant mapping.

The global natures of harmonic mappings are closely related to the curvatures of the manifolds under consideration. For instance, suppose that M and N are compact and that the sectional curvatures of N are nonpositive everywhere. Then we have:

(3) **Uniqueness.** Let $f: M \rightarrow N$ be a harmonic mapping, and assume that there is a point of $f(M)$ where the sectional curvatures of N are negative. Then f is unique in its homotopy class unless $f(M)$ is a closed geodesic γ of N ; and in this case we have uniqueness up to rotation of γ , i.e., an isometry of γ which moves each point of γ a fixed oriented distance along γ .

(4) **Degeneracy.** Suppose further that the \dagger Ricci tensor (R_{ij}) of M is positive semidefinite everywhere. Then the energy density $e(f)$ is a \dagger subharmonic function for every harmonic mapping. This implies that any harmonic mapping $f: M \rightarrow N$ is \dagger totally geodesic. Moreover, if N is of negative sectional curvature, then f is either constant or maps M onto a closed geodesic of N ; if (R_{ij}) is positive definite at some point, then f is constant.

(5) **Finiteness.** Assume now that N is of negative sectional curvature. Then, for each $K \geq 1$, there are only finitely many nonconstant harmonic mappings $f: M \rightarrow N$ of dilatation bounded by K . Here, we say that the dilatation of f is bounded by K if and only if at each point of M we have $df = 0$ or $(\lambda_1/\lambda_2)^{1/2} \leq K$, $\lambda_1 \geq \lambda_2 \geq \dots > 0$ being the positive eigenvalues of the pullback quadratic form $f^*h(x)$ on $T_x(M)$ induced from the metric h of N by f .

D. Harmonic Mappings of a Surface

Let M be a compact surface. Then the energy of a mapping $M \rightarrow N$ is the \dagger Dirichlet-Douglas functional, and harmonic mappings are closely connected with solutions of Plateau's problem (\rightarrow 334 Plateau's Problem). In fact, if a \dagger conformal mapping $M \rightarrow N$ minimizes the \dagger area functional, then it also minimizes the energy functional.

Now let M and N be compact orientable

surfaces whose genera are denoted by p and q , respectively. Then the problem of existence (or nonexistence) of harmonic mappings is well understood. In fact:

(1) When $q \neq 0$, for any metrics g and h on M and N , every homotopy class of mappings $M \rightarrow N$ contains a harmonic mapping.

(2) When $q = 0$ (i.e., N is the 2-sphere S^2), every harmonic mapping whose \dagger degree d satisfies $|d| \geq p$ is holomorphic or antiholomorphic with respect to the complex structures associated with g and h . For example, consider the homotopy classes of mappings from the 2-torus T^2 to S^2 with any metrics. Then all classes with degree $|d| \geq 2$ have harmonic representatives, and any such is holomorphic or antiholomorphic; and the classes with $d = \pm 1$ have no harmonic representatives.

(3) When $q = 0$ and $|d| \leq p - 1$, we have, for every such p and d , a surface M of genus p and a metric h on S^2 such that there exists a harmonic nonholomorphic (and nonantiholomorphic) mapping of degree d from M to S^2 .

E. Existence Theorems

The basic problem in the study of harmonic mappings is to prove their existence in general geometric contexts.

(1) In regard to this problem, translating the problem of the elliptic system $\tau(f) = 0$ into the \dagger initial value problem of the corresponding nonlinear \dagger parabolic system $\partial f / \partial t = \tau(f)$, J. Eells and J. H. Sampson [1] proved that if M and N are compact and if N has nonpositive sectional curvature everywhere, then every homotopy class of mappings $M \rightarrow N$ contains a harmonic mapping that minimizes the energy in that class. Subsequently, the uniqueness of these harmonic mappings was established by P. Hartman [2] in the form stated in Section C.

(2) For harmonic mappings of surfaces, more general existence results have been known.

First, by the \dagger direct method of the calculus of variations, L. Lemaire (1977) and others proved that if M and N are compact and if M is 2-dimensional, then every conjugacy class of homomorphisms $\pi_1(M) \rightarrow \pi_1(N)$ of the fundamental groups is induced by a minimizing harmonic mapping. It follows that if, in particular, the second homotopy group $\pi_2(N)$ of N is zero, then every homotopy class of mappings of a compact surface M to N contains a harmonic representative realizing the minimum of the energy in that class.

On the other hand, by making use of the generalized \dagger Morse theory for a perturbed energy functional, J. Sacks and K. Uhlen-

beck [5] succeeded in giving a satisfactory answer to the structure of $\pi_2(N)$, which is a $\pi_1(N)$ -module, in terms of harmonic mappings. They proved that there exists a generating set for $\pi_2(N)$ consisting of harmonic mappings of spheres that minimize energy and area in their homotopy classes. We note that these harmonic mappings are minimal immersions with \dagger branch points.

(3) Next, we mention the case of harmonic mappings of manifolds with boundary. In this case, we can naturally formulate the \dagger Dirichlet and the \dagger Neumann boundary value problem for harmonic mappings.

In his study of Plateau's problem on Riemannian manifolds, C. Morrey (1948) discussed the Dirichlet problem for harmonic surfaces with boundary.

The problem in arbitrary dimensions has been studied by R. S. Hamilton [3], who extended the result of Eells and Sampson mentioned above to the case where M and N have boundaries. In fact, let M and N be compact Riemannian manifolds with boundary, and assume that N has nonpositive sectional curvature and that the boundary ∂N of N is \dagger convex (or empty). Then there exists a unique minimizing harmonic mapping in each \dagger relative homotopy class determined by the prescribed Dirichlet boundary value. We note that if ∂N is not convex, then it is easy to formulate Dirichlet problems with no solutions. Hamilton also treated the Neumann problem.

Subsequently, S. Hildebrandt, H. Kaul, and K.-O. Widman [4] gave another existence proof of solutions of the Dirichlet problem that covers the case where N admits positive sectional curvature.

References

- [1] J. Eells and J. H. Sampson, Harmonic mappings of Riemannian manifolds, *Amer. J. Math.*, 86 (1964), 109–160.
- [2] P. Hartman, On homotopic harmonic maps, *Canad. J. Math.*, 19 (1967), 673–687.
- [3] R. S. Hamilton, Harmonic maps of manifolds with boundary, *Lecture notes in math.* 471, Springer, 1975.
- [4] S. Hildebrandt, H. Kaul, and K.-O. Widman, An existence theorem for harmonic mappings of Riemannian manifolds, *Acta Math.*, 138 (1977), 1–16.
- [5] J. Sacks and K. Uhlenbeck, The existence of minimal immersions of two-spheres, *Ann. Math.*, (2) 113 (1981), 1–24.
- [6] J. Eells and L. Lemaire, A report on harmonic maps, *Bull. London Math. Soc.*, 10 (1978), 1–68.

- [7] Y.-T. Siu and S.-T. Yau, Compact Kähler manifolds of positive bisectional curvature, *Inventiones Math.*, 59 (1980), 189–204.
- [8] Y.-T. Siu, The complex-analyticity of harmonic maps and the strong rigidity of compact Kähler manifolds, *Ann. Math.*, (2) 112 (1980), 73–111.
- [9] S. Nishikawa and K. Shiga, On the holomorphic equivalence of bounded domains in complete Kähler manifolds of nonpositive curvature, *J. Math. Soc. Japan*, 35 (1983) 273–278.
- [10] T. Ishihara, The index of a holomorphic mapping and the index theorem, *Proc. Amer. Math. Soc.*, 66 (1977), 169–174.
- [11] T. Sunada, Rigidity of certain harmonic mappings, *Inventiones Math.*, 51 (1979), 297–307.

196 (XXI.28) Hilbert, David

David Hilbert (January 23, 1862–February 14, 1943) was born in Königsberg, Germany. He attended the University of Königsberg from 1882 to 1885, when he received his doctoral degree with a thesis on the theory of invariants. It was there that he established a life-long friendship with H. Minkowski. In 1892 he became a professor at the University, and in 1895 he was appointed to a professorship at the University of Göttingen, a position he held until his death. He obtained his basic theorem on invariants between 1890 and 1893, and next began research on the foundations of geometry (→ 155 Foundations of Geometry) and the theory of †algebraic number fields. Concerning the former, he published *Grundlagen der Geometrie* (first edition, 1899), in which he gave the complete axioms of Euclidean geometry and a logical examination of them. Concerning the latter, he systematized all the important known results of algebraic number theory in his monumental *Zahlbericht* (1897). In number theory, he enunciated his significant conjecture on †class field theory. At the international congress of mathematicians held in Paris in 1900, he put forth 23 problems as targets for mathematics of the 20th century (Table 1). Between 1904 and 1906 he conducted research on the †Dirichlet principle of †potential theory and on the direct method in the †calculus of variations. Around 1909 he established the foundations of the theory of †Hilbert spaces. After 1910 he was chiefly involved in research on the †foundations of mathematics, and he advocated the standpoint of †formalism. He is one of the greatest mathematicians of the first half of the 20th century.

Table 1. The 23 Problems of Hilbert

- (1) To prove the continuum hypothesis (→ 33 Axiomatic Set Theory D).
- (2) To investigate the consistency of the axioms of arithmetic (→ 156 Foundations of Mathematics E).
- (3) To show that it is impossible to prove the following fact utilizing only congruence axioms: Two tetrahedra having the same altitude and base area have the same volume. Solved by M. Dehn (1900).
- (4) To investigate geometries in which the line segment between any pair of points gives the shortest path between the pair (→ 155 Foundations of Geometry).
- (5) To obtain the conditions under which a topological group has the structure of a Lie group (→ 423 Topological Groups M). Solved by A. M. Gleason and D. Montgomery and L. Zippin (1952) and H. Yamabe (1953).
- (6) To axiomatize those physical sciences in which mathematics plays an important role.
- (7) To establish the transcendence of certain numbers (→ 430 Transcendental Numbers B). The transcendence of $2\sqrt{2}$, which was one of the numbers put forth by Hilbert, was shown by A. Gel'fond (1934) and T. Schneider (1935).
- (8) To investigate problems concerning the distribution of prime numbers; in particular, to show the correctness of the Riemann hypothesis (→ 450 Zeta Functions). Unsolved.
- (9) To establish a general law of reciprocity (→ 59 Class Field Theory A). Solved by T. Takagi (1921) and E. Artin (1927).
- (10) To establish effective methods to determine the solvability of Diophantine equations (→ 97 Decision Problem; 182 Geometry of Numbers). Solved affirmatively for equations of two unknowns by A. Baker, *Philos. Trans. Roy. Soc. London*, (A) 263 (1968); solved negatively for the general case by Yu. V. Matiyasevich (1970).
- (11) To investigate the theory of quadratic forms over an arbitrary algebraic number field of finite degree (→ 348 Quadratic Forms).
- (12) To construct class fields of algebraic number fields (→ 73 Complex Multiplication).
- (13) To show the impossibility of the solution of the general algebraic equation of the seventh degree by compositions of continuous functions of two variables. Solved negatively. In general, V. I. Arnold proved that every real, continuous function $f(x_1, x_2, x_3)$ on $[0, 1]$ can be represented in the form $\sum_{i=1}^9 h_i(g_i(x_1, x_2, x_3))$, where h_i and g_i are real, continuous functions, and A. N. Kolmogorov proved that $f(x_1, x_2, x_3)$ can be represented

in the form $\sum_{i=1}^7 h_i(g_{i1}(x_1) + g_{i2}(x_2) + g_{i3}(x_3))$, where h_i and g_{ij} are real, continuous functions and g_{ij} can be chosen once for all independently of f (*Dokl. Akad. Nauk SSSR*, 114 (1957), *Amer. Math. Soc. Transl.*, 28 (1963)).

(14) Let k be a field, x_1, \dots, x_n be variables, and $f_i(x_1, \dots, x_n)$ be given polynomials in $k[x_1, \dots, x_n]$ ($i = 1, \dots, m$). Furthermore, let R be the ring formed by rational functions $F(X_1, \dots, X_m)$ in $k(X_1, \dots, X_m)$ such that $F(f_1, \dots, f_m) \in k[x_1, \dots, x_n]$. The problem is to determine whether the ring R has a finite set of generators. Solved negatively by M. Nagata, *Amer. J. Math.*, 81 (1959).

(15) To establish the foundations of algebraic geometry (\rightarrow 12 Algebraic Geometry). Solved by B. L. van der Waerden (1938–1940), A. Weil (1950), and others.

(16) To conduct topological studies of algebraic curves and surfaces.

(17) Let $f(x_1, \dots, x_n)$ be a rational function with real coefficients that takes a positive value for any real n -tuple (x_1, \dots, x_n) . The problem is to determine whether the function f can be written as the sum of squares of rational functions (\rightarrow 149 Fields O). Solved in the affirmative by E. Artin (1927).

(18) To express Euclidean n -space as a disjoint union $\bigcup_{\lambda} P_{\lambda}$, where each P_{λ} is congruent to one of a set of given polyhedra.

(19) To determine whether the solutions of regular problems in the calculus of variations are necessarily analytic (\rightarrow 323 Partial Differential Equations of Elliptic Type). Solved by S. N. Bernshtein, I. G. Petrovskii, and others.

(20) To investigate the general boundary value problem (\rightarrow 120 Dirichlet Problem; 323 Partial Differential Equations of Elliptic Type).

(21) To show that there always exists a linear differential equation of the Fuchsian class with given singular points and monodromic group (\rightarrow 253 Linear Ordinary Differential Equations (Global Theory)). Solved by H. Röhrl and others (1957).

(22) To uniformize complex analytic functions by means of automorphic functions (\rightarrow 367 Riemann Surfaces). Solved for the case of one variable by P. Koebe (1907).

(23) To develop the methodology of the calculus of variations (\rightarrow 46 Calculus of Variations).

References

- [1] D. Hilbert, *Gesammelte Abhandlungen I–III*, Springer, 1932–1935 (Chelsea, 1967).
 [2] D. Hilbert, *Grundlagen der Geometrie*, Teubner, seventh edition, 1930.

[3] D. Hilbert, *Grundzüge der allgemeinen Theorie der linearen Integralgleichungen*, Teubner, second edition, 1924 (Chelsea, 1953).

[4] D. Hilbert and W. Ackermann, *Grundzüge der theoretischen Logik*, Springer, third edition, 1949; English translation, *Principles of mathematical logic*, Chelsea, 1950.

[5] D. Hilbert and P. Bernays, *Grundlagen der Mathematik*, Springer, second edition, I, 1968; II, 1970.

[6] F. Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert I*, Springer, 1926 (Chelsea, 1956).

[7] H. Weyl, David Hilbert and his mathematical work, *Bull. Amer. Math. Soc.*, 50 (1944), 612–654.

[8] C. Reid, Hilbert. With an appreciation of Hilbert's mathematical work by H. Weyl, Springer, 1970.

197 (XII.2) Hilbert Spaces

A. General Remarks

The theory of Hilbert spaces arose from problems in the theory of * integral equations. D. Hilbert noticed that a linear integral equation can be transformed into an infinite system of linear equations for the * Fourier coefficients of the unknown function. He considered the linear space l_2 consisting of all sequences of numbers $\{x_n\}$ for which $\sum_{n=1}^{\infty} |x_n|^2$ is finite, and defined for each pair of elements $x = \{x_n\}$, $y = \{y_n\} \in l_2$ their inner product as $(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n$. The space l_2 can be regarded as an infinite-dimensional extension of the notion of a Euclidean space. F. Riesz considered the space of functions now termed L_2 -space and succeeded in giving a satisfactory answer to the Fourier expansion problem. In his book [3], J. von Neumann established a rigorous foundation of quantum mechanics employing Hilbert spaces and the spectral expansion of self-adjoint operators. The following axiomatic definition (\rightarrow Section B) of Hilbert spaces is due to von Neumann. H. Weyl later justified the * Dirichlet principle of Riemann by the method of orthogonal projection in a Hilbert space, and thus paved the way for the function-analytic study of differential equations.

B. Definition

Let K be the field of complex or real numbers, the elements of which we denote by α, β, \dots

Let H be a †linear space over K , and to any pair of elements $x, y \in H$ let there correspond a number $(x, y) \in K$ satisfying the following five conditions: (i) $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$; (ii) $(\alpha x, y) = \alpha(x, y)$; (iii) $(x, y) = \overline{(y, x)}$; (iv) $(x, x) \geq 0$; and (v) $(x, x) = 0 \Leftrightarrow x = 0$. Then we call H a **pre-Hilbert space** and (x, y) the **inner product** of x and y .

With the norm $\|x\| = \sqrt{(x, x)}$, H is a †normed linear space. If H is †complete with respect to the distance $\|x - y\|$ (i.e., $\|x_n - x_m\| \rightarrow 0$ ($m, n \rightarrow \infty$)) implies the existence of $\lim x_n = x$), then we call H a **Hilbert space**. According as K is complex or real, we call H a **complex** or **real Hilbert space**. A Hilbert space is a †Banach space.

A normed linear space with norm $\|x\|$ can be made a pre-Hilbert space, by defining an inner product (x, y) so that $\|x\| = \sqrt{(x, x)}$, if and only if the equality $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ holds for any x, y .

C. Orthonormal Sets

Two elements $x, y \in H$ are said to be mutually **orthogonal** if $(x, y) = 0$. A subset Σ of H is called an **orthogonal set** (or **system**) if $0 \notin \Sigma$ and every distinct pair $x, y \in \Sigma$ is mutually orthogonal. If every element of an orthogonal set Σ is of norm 1, then Σ is called an **orthonormal set**. Any orthogonal set $\Sigma = \{x_i\}$ can be **normalized** into an orthonormal set $\{x_i/\|x_i\|\}$. A maximal orthonormal set is called a **complete orthonormal set** or an **orthonormal basis**. All the complete orthonormal sets of a given H have the same cardinal number, which we call the **dimension** of H . Two Hilbert spaces are isomorphic if and only if they have the same dimension.

Let $\Sigma = \{x_i\}$ be an orthonormal set. Then for every $x \in H$, its **Fourier coefficients** (x, x_i) vanish for all but a countable number of i , and the **Bessel inequality** $\|x\|^2 \geq \sum_i |(x, x_i)|^2$ holds. The following three statements are equivalent in a Hilbert space: (i) Σ is complete; (ii) **Parseval's equality** $\|x\|^2 = \sum_i |(x, x_i)|^2$ holds for every x ; (iii) every x can be expanded in a **Fourier series** $x = \sum_i (x, x_i)x_i$ (\rightarrow 317 Orthogonal Functions).

D. Examples of Hilbert Spaces

The space l_2 (\rightarrow Section A) is a Hilbert space of dimension \aleph_0 . The †function space L_2 on a measure space (X, μ) is a Hilbert space if the inner product of $f, g \in L_2$ is defined by $(f, g) = \int_X f \bar{g} d\mu$. In the case of the †Lebesgue measure in a Euclidean space, L_2 is of dimension \aleph_0 , so that it is a Hilbert space isomorphic to l_2 . Further examples of Hilbert spaces are

$A_2(\Omega)$, $W_2^1(\Omega)$ ($= H^1(\Omega)$), and $H_0^1(\Omega)$ (\rightarrow 168 Function Spaces).

E. Closed Linear Subspaces and Projections

Let M be a **closed linear subspace** of a Hilbert space H , i.e., a linear subspace that is closed in the norm topology of H . It is a Hilbert space with respect to the restriction of the inner product in H . For a given M the set of all $x \in H$ such that $(x, y) = 0$ for every $y \in M$ forms a closed linear subspace M^\perp called the **orthogonal complement** of M . The orthogonal complement of M^\perp is M (i.e., $M^{\perp\perp} = M$), and H is the **direct sum** of M and M^\perp (i.e., every $x \in H$ can be uniquely represented as $x = y + z$, $y \in M$, $z \in M^\perp$, and $\|x\|^2 = \|y\|^2 + \|z\|^2$). Thus the quotient space H/M is isomorphic to M^\perp and is also a Hilbert space. The operator P_M that maps x to y is called the **projection** (or **orthogonal projection** or **projection operator**) to M . A bounded linear operator P is a projection if and only if it is idempotent ($P^2 = P$) and self-adjoint ($((Px, y) = (x, Py)$ for any $x, y \in H$) (\rightarrow 251 Linear Operators).

F. Conjugate Spaces

A linear operator from H to K is called a **linear functional**. The set H' of all continuous linear functionals f on H forms a Hilbert space with norm $\|f\| = \sup\{|f(x)| : \|x\| = 1\}$. For every $f \in H'$ there exists a unique $y \in H$ such that $f(x) = (x, y)$ for all $x \in H$ (**Riesz's theorem**), and the correspondence $f \rightarrow y$ gives an †antilinear isometric operator from H' onto H (for †linear operators on Hilbert spaces \rightarrow 68 Compact and Nuclear Operators; 251 Linear Operators; 390 Spectral Analysis of Operators).

References

- [1] D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Teubner, second edition, 1924 (Chelsea, 1953).
- [2] E. Hellinger and O. Toeplitz, *Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten*, Enzykl. Math., Teubner, 1928 (Chelsea, 1953).
- [3] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, 1932.
- [4] J. von Neumann, *Collected works II, III*, Pergamon, 1961.
- [5] M. H. Stone, *Linear transformations in Hilbert space and their applications to analysis*, Amer. Math. Soc. Colloq. Publ., 1932.

- [6] B. Sz.-Nagy, *Spektraldarstellung linearer Transformationen des Hilbertschen Raumes*, Erg. Math., Springer, 1942.
- [7] F. Riesz and B. Sz.-Nagy, *Leçons d'analyse fonctionnelle*, Akademiai Kiadó, third edition, 1955; English translation, *Functional analysis*, Ungar, 1955.
- [8] N. I. Akhiezer and I. M. Glazman, *Theory of linear operators in Hilbert space*, Ungar, I, 1961; II, 1963. (Original in Russian, 1950.)
- [9] K. Yosida, *Functional analysis*, Springer, 1965.
- [10] R. Courant and D. Hilbert, *Methods of mathematical physics I*, Interscience, 1953.
- [11] P. R. Halmos, *Introduction to Hilbert space and the theory of spectral multiplicity*, Chelsea, second edition, 1957.
- [12] P. R. Halmos, *A Hilbert space problem book*, Van Nostrand, 1967.

198 (XI.1) Holomorphic Functions

A. Differentiation of Complex Functions

Let $f(z)$ be a * complex-valued function defined in an open set D in the * complex plane \mathbb{C} . We say that $f(z)$ is **differentiable** at z if the limit

$$\lim_{h \rightarrow 0} (f(z+h) - f(z))/h = f'(z) \quad (1)$$

exists and is finite as the complex number h tends to zero. We call $f'(z)$ the **derivative** of $f(z)$ at z . This definition is a formal extension of the definition of differentiability of a function of a real variable to that of a complex variable (\rightarrow 106 Differential Calculus), but it is a much stronger condition than the differentiability of a real function, since $z+h$ in (1) may be an arbitrary point in a 2-dimensional neighborhood of z . Hence many results essentially different from those for functions of a real variable follow from it.

If a function $f(z)$ is differentiable at each point of an open set D , it is said to be **holomorphic** (or **regular**) in D , or $f(z)$ is a **holomorphic function** on D . (For the definition of holomorphy of a complex-valued function of several complex variables \rightarrow 21 Analytic Functions of Several Complex Variables C.)

Let E be an arbitrary nonempty subset of \mathbb{C} . We say that $f(z)$ is holomorphic on E if it is defined in an open set D containing E and is holomorphic on D . Some results valid for differentiable real functions also hold for holomorphic functions. For instance, the derivative of a sum, product, or quotient is given by the usual rules. The derivative of a composite function is determined by the chain rule.

The set of all functions holomorphic in a * domain D forms a * ring.

Suppose $f(z)$ is holomorphic on D and $f'(z_0) \neq 0$, $z_0 \in D$. Then two curves that form an angle at z_0 are mapped by f to two curves forming the same angle at $f(z_0)$. Because of this property, the mapping f is said to be **conformal** at all points z with $f'(z) \neq 0$.

The following four conditions are equivalent for a function $f = u + iv$ defined on an open set D . (1) f is holomorphic in D . (2) $u = u(x, y)$ and $v = v(x, y)$ are * totally differentiable at each point $z = x + iy$ and satisfy the **Cauchy-Riemann differential equations**

$$\partial u / \partial x = \partial v / \partial y, \quad \partial u / \partial y = -\partial v / \partial x.$$

(3) f is represented by a * power series $\sum_{n=0}^{\infty} c_n(z - a)^n$ in a neighborhood of each point a of D ; that is, $f(z)$ is **analytic** in D . (4) f is continuous and $\int_C f(z) dz = 0$ for every rectifiable Jordan closed curve C whose interior is contained, together with C , in D . The proposition that (1) implies (4) is called **Cauchy's integral theorem**, and the proposition that (4) implies (1) is called **Morera's theorem**.

The hypothesis of Morera's theorem can be weakened as follows: Let $f(z)$ be continuous in a domain D . If $\int_C f(z) dz = 0$ for every rectangle C in D with sides parallel to the axes and whose interior consists of only points of D , then $f(z)$ is holomorphic in D . In the statement of this theorem, if we let C be an arbitrary circle, we get the same conclusion.

The following complex differential operators are often useful:

$$\frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right), \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right).$$

Generally, $(\partial \bar{\phi} / \partial \bar{z}) = \partial \bar{\phi} / \partial \bar{z}$. The Cauchy-Riemann equations above can be expressed in a single equation: $\partial f / \partial \bar{z} = 0$. If f is holomorphic, $\partial f / \partial \bar{z} = f'(z)$.

In order to show that $f = u + iv$ is holomorphic in D , assumption (2) can be weakened. Actually, we have the **Looman-Men'shov theorem**: Suppose that u and v are continuous in D , $\partial u / \partial x$, $\partial u / \partial y$, $\partial v / \partial x$, and $\partial v / \partial y$ exist at every point of D except for at most a countable number of points, and the Cauchy-Riemann equations hold in D except for a set of 2-dimensional * measure zero; then $f = u + iv$ is holomorphic in D . D. E. Men'shov extended this theorem and obtained various conditions for holomorphy. For example, he proved the following theorem: If f is a topological mapping of D and f is conformal in D (i.e.,

$$\lim_{h \rightarrow 0} \arg(f(z+h) - f(z))/h$$

exists) except for at most a countable number of points, then f is holomorphic in D .

As another type of sufficient condition for holomorphy, we have the proposition: If f is locally † Lebesgue integrable in D and satisfies the Cauchy-Riemann equation $f_{\bar{z}} = 0$ in the sense of † distribution, then there exists a holomorphic function g in D such that $g = f$ † almost everywhere.

B. Cauchy's Integral Theorem

Cauchy's integral theorem can be stated as follows: If $f(z)$ is a holomorphic function in a † simply connected domain D in the complex plane, the equality $\int_C f(z) dz = 0$ holds for every (rectifiable) closed curve C in D . In particular, the integral $\int_{\alpha}^{\beta} f(z) dz$ ($\alpha, \beta \in D$) is uniquely determined by α and β provided that its path of integration lies in D . The function $F(z) = \int_{\alpha}^z f(\zeta) d\zeta$ ($z \in D$) is called the **indefinite integral** of f . $F(z)$ is holomorphic in D and $F'(z) = f(z)$.

In the proof of his integral theorem, Cauchy assumed the existence and continuity of the derivative $f'(z)$ in D . However, E. Goursat proved the theorem utilizing only the existence of $f'(z)$. Actually, by virtue of the integral formula (2) in this section, the existence of $f'(z)$ in D implies the continuity of $f'(z)$. This is sometimes called **Goursat's theorem**.

Let C, C_1, C_2, \dots, C_n be rectifiable Jordan curves. Suppose that C_1, C_2, \dots, C_n are in the interior of C and that each one lies in the exterior of the others. If $f(z)$ is holomorphic in the region D bounded by these $n+1$ curves and continuous on $D \cup C \cup C_1 \cup \dots \cup C_n = \bar{D}$, then we have

$$\int_C f(z) dz = \sum_{k=1}^n \int_{C_k} f(z) dz.$$

Here the curvilinear integrals are taken in the **positive direction** (i.e., we take the direction such that $\int_C (z-a)^{-1} dz = \int_{C_k} (z-a)^{-1} dz = 2\pi i$ for a point a in the interior of C or C_k , respectively). Henceforth, an integral along a closed curve is taken in the positive direction unless otherwise noted. Cauchy's integral theorem under the assumption that f is holomorphic in D and is continuous on \bar{D} is sometimes called the **stronger form of Cauchy's integral theorem**.

Under the same assumptions as in the stronger form, we have **Cauchy's integral formula** for $z \in D$:

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z} d\zeta - \frac{1}{2\pi i} \int_{C_1} \frac{f(\zeta)}{\zeta - z} d\zeta \dots - \frac{1}{2\pi i} \int_{C_n} \frac{f(\zeta)}{\zeta - z} d\zeta. \quad (2)$$

This integral formula expresses the value of

$f(z)$ at a point z in the domain D in terms of the values of f on the boundary of D . In particular, when $n=0$, the integral formula reads as

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z} d\zeta. \quad (3)$$

Furthermore, if C is a circle $|z| = R$ (i.e., D is the disk $|z| < R$), we obtain **Poisson's integral formula**:

$$\begin{aligned} f(z) &= \\ \frac{1}{2\pi} \int_0^{2\pi} f(Re^{i\varphi}) \frac{R^2 - r^2}{R^2 + r^2 - 2Rr \cos(\theta - \varphi)} d\varphi, \\ f(z) &= \\ \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re} f(Re^{i\varphi}) \frac{Re^{i\varphi} + z}{Re^{i\varphi} - z} d\varphi + i \operatorname{Im} f(0), \\ z &= re^{i\theta}, \quad 0 \leq r < R. \end{aligned} \quad (3')$$

Formula (3)' is valid for a † harmonic function.

Let C be a rectifiable curve and $f(\zeta)$ be a continuous function defined on C ; then the **integral of Cauchy type**

$$F(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z} d\zeta$$

is holomorphic outside C . The n th derivative $F^{(n)}$ of F is given by $(n!/2\pi i) \int_C f(\zeta)/(\zeta - z)^{n+1} d\zeta$; moreover, F can be expanded in a Taylor series about every point a outside C :

$$F(z) = \sum_{n=0}^{\infty} a_n (z-a)^n, \quad a_n = \frac{F^{(n)}(a)}{n!},$$

which converges in $|z-a| < \rho$, ρ being the distance from a to C . In particular, formula (3) implies that a holomorphic function f is infinitely differentiable and is expanded in a Taylor series about every point of D as above.

Let C be a closed curve not passing through a point a . Then the integral $(1/2\pi i) \int_C dz/(z-a)$ is an integer. It is called the **winding number** of C about a and is denoted by $n(C; a)$. A cycle γ (a finite sum of oriented closed curves) in an open set D is said to be **homologous to zero** in D if $n(\gamma; a) = 0$ for all points a in the complement of \bar{D} . The general form of Cauchy's integral theorem is stated as follows. If f is holomorphic in D , then $\int_{\gamma} f(z) dz = 0$ for every cycle γ which is homologous to zero in D (E. Artin). From this we have the general form of Cauchy's integral formula: if f is holomorphic in a domain D , then

$$n(\gamma; z) \cdot f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{\zeta - z} dz, \quad z \in D - \gamma,$$

for every cycle γ which is homologous to zero in D .

C. Zero Points

Let $f(z)$ be a holomorphic function not identically equal to zero. If $f(a) = 0$, we call a a **zero point** of f . Every zero point of f is an isolated point, and there exists a unique positive integer k and a function h holomorphic at a such that

$$f(z) = (z-a)^k g(z), \quad g(a) \neq 0. \quad (4)$$

We call k the **order of the zero point** a and a a **zero point of the k th order**. The equality (4) implies that the † Taylor series of $f(z)$ at a begins with the term $c_k(z-a)^k$. Suppose that a is a zero point of $f(z) - \gamma$ of the k th order; then we call a a **γ -point of the k th order**.

For a function $f(z)$ defined in a neighborhood of the † point at infinity, we set $f(1/w) = g(w)$ ($f(\infty) = g(0)$) and call f holomorphic at ∞ if g is holomorphic at 0; f is said to have a zero of **order k** at ∞ if g has a zero of order k at 0. If two functions f and g are holomorphic in D and $f(z) = g(z)$ on a subset E that has an † accumulation point in D , then f is identically equal to g in D (**theorem of identity** or **uniqueness theorem**) since the zeros of holomorphic functions must be isolated.

D. Isolated Singularities

Let $f(z)$ be holomorphic in an annulus $D = \{z \mid R_1 < |z-a| < R_2, 0 \leq R_1 < R_2 \leq +\infty\}$. Then $f(z)$ is expanded in the † Laurent series

$$f(z) = \sum_{n=-\infty}^{+\infty} c_n(z-a)^n. \quad (5)$$

This is called the **Laurent expansion** of f about a . The coefficients c_n are given by $c_n = (1/2\pi i) \int_C f(\zeta) d\zeta / (\zeta-a)^{n+1}$ with $C = \{z \mid |z-a| = r\}$, $R_1 < r < R_2$. In particular, if $f(z)$ is holomorphic in $D = \{z \mid 0 < |z-a| < R\}$ (or, if $a = \infty$, in $D = \{z \mid R < |z| < +\infty\}$) but not holomorphic in $D \cup \{a\}$, we call a an **isolated singular point** (or **isolated singularity**) of f . By utilizing the † local canonical parameter $t = z-a$ (or $t = 1/z$ for $a = \infty$), the Laurent expansion (5) of f is then written as $f(z) = \sum_{n=-\infty}^{-1} c_n t^n + \sum_{n=0}^{+\infty} c_n t^n$. The second sum is an ordinary power series, called the **holomorphic part** of $f(z)$. The first sum is a power series of $1/t$ with no constant term, called the **singular part** of $f(z)$ at a or the **principal part** of the singularity (or of the Laurent expansion at a).

If we have $\lim_{t \rightarrow 0} t f(z) = 0$, the Laurent expansion (5) of $f(z)$ lacks its singular part, and the limit of $f(z)$ exists as $t \rightarrow 0$ ($z \rightarrow a$) and is equal to c_0 . If we set $f(a) = c_0$, then the function $f(z)$ is holomorphic in $D \cup \{a\}$. In this case, the point a is called a **removable singular-**

ity. If $f(z)$ is bounded in a neighborhood of a singularity a , then a is removable (**Riemann's theorem**). Usually, we assume that the removable singularities of a function have already been removed in this way.

When the singular part of $f(z)$ at a exists and consists of a finite number of terms, the point a is called a **pole**; when it consists of an infinite number of terms, the point is called an **essential singularity**. If a is a pole, $f(z)$ is represented by the Laurent series $\sum_{n=-k}^{\infty} c_n t^n$ ($c_{-k} \neq 0$) and $f'(z) \rightarrow \infty$ as $z \rightarrow a$. In this case, the index k is called the **order of the pole** a . Then a relation such as (4) holds, where the index k is replaced by $-k$. Hence the point a is sometimes called a **zero point of the $-k$ th order**. If a is an essential singularity, then for an arbitrary number c there exists a sequence z_n converging to a such that $\lim_{n \rightarrow \infty} f(z_n) = c$ (the **Casorati-Weierstrass theorem** or simply **Weierstrass's theorem**). Related to Weierstrass's theorem, we have † Picard's theorem, which gives a detailed description of the behavior of a function around its singularities.

E. Residues

Let $a (\neq \infty)$ be an isolated singularity of $f(z)$. Then the coefficient c_{-1} of $(z-a)^{-1}$ in the Laurent expansion (5) of $f(z)$ is called the **residue** of $f(z)$ at a and is denoted by $\text{Res}[f]_a$, $R(a; f)$, or $R(a)$ if we need not indicate f . We have

$$R(a) = c_{-1} = \frac{1}{2\pi i} \int_{|\zeta-a|=r} f(\zeta) d\zeta,$$

where the integral is taken in the positive direction along a path for $0 < r < R$. If $f(z)$ is holomorphic at $z = a$, then $R(a) = 0$. If $f(z)$ has a pole of the first order at a ,

$$R(a) = \lim_{z \rightarrow a} (z-a)f(z).$$

The residue at the point at infinity is defined to be $-a_{-1}$, where a_{-1} is the coefficient of $1/z$ of the Laurent expansion of $f(z)$ at ∞ : $f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$, and we have

$$-a_{-1} = \frac{-1}{2\pi i} \int_{|\zeta|=r} f(\zeta) d\zeta, \quad R^{-1} < r < +\infty.$$

Thus the notion of the residue of $f(z)$ is actually related to the differential form $f(z)dz$ and not to $f(z)$ itself.

From the first formula in this section and the formula for $-a_{-1}$, the **residue theorem** follows (Cauchy, 1825): Let C be a rectifiable Jordan curve in the complex plane. Let a_1, \dots, a_m be a finite number of points inside C , and let D be a domain containing C and its

interior. If $f(z)$ is a function holomorphic in $D - \{a_1, \dots, a_m\}$, we have

$$\frac{1}{2\pi i} \int_C f(z) dz = \sum_{n=1}^m R(a_n).$$

Furthermore, if $f(z)$ is holomorphic in the extended complex plane (including the point at infinity) except for a finite number of poles, the sum of all residues is equal to zero.

F. Calculus of Residues

The **calculus of residues** is a field of calculus based on application of the notion of residues. For example, we have methods for the calculation of definite integrals. Actually, one of the reasons why Cauchy studied the theory of complex functions was that he believed that the theory would provide a unified method of computing definite integrals. For example, if $\varphi(z)$ is a rational function without poles on the real axis and with a zero point at infinity whose order is at least 2, then we have

$$\int_{-\infty}^{\infty} \varphi(x) dx = 2\pi i \sum_{\text{Im } \alpha > 0} R(\alpha; \varphi(z)), \quad (6)$$

$$\int_{-\infty}^{\infty} e^{ix} \varphi(x) dx = 2\pi i \sum_{\text{Im } \alpha > 0} R(\alpha; e^{iz} \varphi(z)). \quad (7)$$

Here the sums are taken over all the poles in the upper half-plane. Formula (7) is valid also for a rational function $\varphi(z)$ with a simple zero at infinity. If $\varphi(z)$ has simple poles at a_k ($k = 1, \dots, n$) on the real axis, then we take the principal values of the integrals at those poles and add $\pi i R(a_k)$ ($k = 1, \dots, n$) to the terms on the right-hand side of (6) and (7). Sometimes we use the residue theorem to obtain the value of the sum of a series (e.g., the †Gaussian sum) by expressing it as an integral.

Let $f(z)$ be a single-valued function that is †meromorphic and not identically equal to zero in a domain D , and let $\varphi(z)$ be a function holomorphic in D . Draw a rectifiable Jordan curve C in D such that the interior of C is contained in D and $f(z)$ has neither zeros nor poles on C . Let $\alpha_1, \dots, \alpha_N$ and β_1, \dots, β_P be the zeros and poles inside C , respectively (where each of them is repeated as often as its order). Then we have

$$\sum_{n=1}^N \varphi(\alpha_n) - \sum_{p=1}^P \varphi(\beta_p) = \frac{1}{2\pi i} \int_C \varphi(z) \frac{f'(z)}{f(z)} dz.$$

If $\varphi(z) = 1$, we get

$$\frac{1}{2\pi} \int_C d \arg f(z) = N - P.$$

This is called the **argument principle**. Next, let $f(z)$ be a function meromorphic for $|z| <$

$R \leq +\infty$, $f(0) \neq 0, \neq \infty$, and set $\varphi(z) = \log z$. Take C as a closed curve consisting of the boundary of an annulus $0 < \rho < |z| < r < R$ (where ρ is sufficiently small) and two sides of a suitable †crosscut joining a point of $|z| = \rho$ and $|z| = r$. Then we obtain **Jensen's formula**:

$$\log \left| \frac{\alpha_1, \alpha_2, \dots, \alpha_N}{\beta_1, \beta_2, \dots, \beta_P} \right| = \log |f(0)| + (N - P) \log r - \frac{1}{2\pi} \int_0^{2\pi} \log |f(re^{i\psi})| d\psi.$$

The argument principle can be utilized to prove **Rouché's theorem**: Let $f(z)$ and $g(z)$ be functions holomorphic in a domain D that contains a rectifiable Jordan curve C and its interior. Suppose that $f(z) + \lambda g(z)$ never vanishes on C for any λ with $0 \leq \lambda \leq 1$. Then the number of zeros of $f(z)$ in the interior of C is equal to that of $f(z) + g(z)$. If $|f(z)| > |g(z)|$ on C or $\arg f(z) - \arg g(z) \neq (2n + 1)\pi$ (n is an integer), the hypothesis of Rouché's theorem is satisfied. This theorem is useful in proving the existence of a zero of a complex function (for example, a polynomial) and in finding its position.

G. Analytic Continuations

Let $f(z)$ be a holomorphic function in a domain D of the complex plane \mathbb{C} and D^* be a domain containing D as a proper subset. If there exists a function $F(z)$ holomorphic in D^* that coincides with $f(z)$ in D , then $F(z)$ is called an **analytic continuation** (or **analytic prolongation**) of $f(z)$ from D to D^* . By the theorem of identity an analytic continuation of $F(z)$ is uniquely determined if it exists.

The function $f_1(z)$ defined by the power series $P(z; a) = \sum_{n=0}^{\infty} a_n(z - a)^n$ with the radius of convergence $r_1 > 0$ is holomorphic in the domain $D_1: |z - a| < r_1$, and at a point b of D_1 it can be expanded into a power series $P(z; b) = \sum_{n=0}^{\infty} b_n(z - b)^n$ with the radius of convergence $r_2 (\geq r_1 - |b - a|)$. If $r_2 > r_1 - |b - a|$, the domain $D_2: |z - b| < r_2$ is not entirely contained in D_1 . Let $f_2(z)$ be the function defined in D_2 by $P(z; b)$. Then the function $F(z)$ that is equal to $f_1(z)$ in D_1 and to $f_2(z)$ in D_2 is an analytic continuation of $f_1(z)$ from D_1 to $D_1 \cup D_2$ (a **direct analytic continuation** by power series).

We have the following classical theorems about analytic continuations:

Let D_1 and D_2 be two disjoint domains, and suppose that their respective boundaries C_1 and C_2 are †rectifiable simple closed curves and that the intersection of C_1 and C_2 contains an open arc Γ . If two holomorphic functions $f_1(z)$, $f_2(z)$ defined in D_1 and D_2 , respectively, have finite common †boundary values at every

point of Γ , then there exists an analytic continuation $F(z)$ of $f_1(z)$ and $f_2(z)$ to $D_1 \cup \Gamma \cup D_2$ (**Painlevé's theorem**). We sometimes call $f_2(z)$ a continuation of $f_1(z)$ beyond Γ . If Γ is not rectifiable, the continuation beyond Γ does not exist, in general.

Let $f(z)$ be holomorphic in a \dagger Jordan domain D lying in the half-plane on one side of the real axis and containing an open interval I of the real axis in its boundary. If $f(z)$ has finite real boundary values at every point of I , then it can be continued analytically beyond I to the other side of the real axis; there the continued function is given by $\overline{f(\overline{z})}$ (**Schwarz's principle of reflection**). This theorem can be generalized to the case where the real interval is replaced by an \dagger analytic curve.

A **harmonic continuation** of \dagger harmonic functions is defined analogously to analytic continuation. Let D be a Jordan domain lying in the half-plane on one side of the real axis and having an open interval I on the real axis as a part of its boundary. If $u(z)$ is harmonic in D and has the boundary value 0 at every point of I , then $u(z)$ has a harmonic continuation beyond I .

For other properties of holomorphic functions \rightarrow 43 Bounded Functions; 429 Transcendental Entire Functions.

H. Analytic Functions

A real-valued function $f(t)$ of a real variable t is said to be **analytic** at $t = t_0$ if it can be represented by a \dagger power series in $t - t_0$ in a neighborhood of t_0 in \mathbf{R} . If $f(t)$ is defined on an open set of \mathbf{R} at every point of which it is analytic, then $f(t)$ is called an **analytic function**, or, more precisely, a **real analytic function**.

Analogously, a complex-valued function $f(z)$ of a complex variable z defined on a \dagger domain D of the complex plane \mathbf{C} is said to be **analytic** at $z = z_0 (\in D)$ if it can be represented by a power series in $z - z_0$ in a neighborhood of z_0 in \mathbf{C} , and $f(z)$ is an **analytic function** in D if it is analytic at every point of D . In the remainder of this article, we are concerned with analytic functions in this sense. To distinguish them from the real case, they are also called **complex analytic functions**. A complex analytic function $f(z)$ defined on D is \dagger differentiable in D ; therefore it is \dagger holomorphic in D . The converse is also true. Thus the term "analytic function" is synonymous with "holomorphic function" insofar as it concerns a complex function (i.e., a complex-valued function of a complex variable) on a domain, but in the theory of functions it takes on an additional meaning that is explained in the following section.

I. Analytic Functions in the Sense of Weierstrass

Let a be a point of the \dagger z -sphere and t the \dagger local canonical parameter at a ; i.e., $t = z - a$ if $a \neq \infty$ and $t = z^{-1}$ if $a = \infty$. If a power series $P(z; a) = \sum_{n=0}^{\infty} c_n t^n$ has a positive radius r_a of convergence, we call $P(z; a)$ a **function element** with center a on the z -sphere, after K. Weierstrass. $P(z; a) = \sum_{n=0}^{\infty} c_n (z - a)^n$ if $a \neq \infty$, and $P(z; a) = \sum_{n=0}^{\infty} c_n z^{-n}$ if $a = \infty$. These represent a holomorphic function in $|z - a| < r_a$ or in $r_a^{-1} < |z| \leq \infty$, respectively. If b is a point inside the circle of convergence of the function element $P(z; a)$, by the \dagger Taylor expansion of $P(z; a)$ at $z = b$, we obtain the power series $P(z; b)$ in $z - b$, which is a direct analytic continuation of $P(z; a)$. Let a and b be two points on the z -sphere, and let $C: z = z(s)$ ($0 \leq s \leq 1$, $z(0) = a$, $z(1) = b$) be a curve joining a and b . We say that $P(z; a)$ is **analytically continuable** along C and that we obtain $P(z; b)$ at the end point b by the **analytic continuation** of $P(z; a)$ along C if the following two conditions are satisfied: (i) To every $s \in [0, 1]$ there corresponds a function element $P(z; z(s))$ with center $z(s)$; (ii) for every $s_0 \in [0, 1]$, we can take a suitable subarc $z = z(s)$ ($|s - s_0| \leq \varepsilon$, $\varepsilon > 0$) of C contained inside the circle of convergence of $P(z; z(s_0))$ such that every function element $P(z; z(s))$ with $|s - s_0| \leq \varepsilon$ is a direct analytic continuation of $P(z; z(s_0))$. When $P(z; a)$ and the curve C are given, the analytic continuation along C is uniquely determined (**uniqueness theorem of the analytic continuation**).

Given a function element $P(z; a)$ with center a , the set of all function elements obtained by every possible analytic continuation along every curve starting from a is called an **analytic function in the sense of Weierstrass** determined by $P(z; a)$. In this definition, we can restrict the curves to polygonal lines. An analytic function in this sense is completely determined by a single arbitrary function element belonging to it, so two analytic functions are identically equal if they have a common function element.

A \dagger germ of a holomorphic function is identical to a function element, and the set of all germs has the natural structure of a \dagger sheaf \mathcal{O} . In the terminology of sheaves, an analytic function is a connected component of \mathcal{O} , and an **analytic continuation along a curve C** is a continuous curve Γ in \mathcal{O} whose projection is C .

J. Values and Branches of Analytic Functions

The value of an analytic function at a point b is, by definition, the value at b of its function elements with center b (whose existence is

assumed; there may be several such elements). An analytic function is, in general, a multiple-valued function because analytic continuations along different curves with the same end points may lead to different function elements. For a given analytic function $f(z)$, if the maximal number of its function elements with the same center is n , we say it is n -valued, and if $n \geq 2$ we say it is **multiple-valued** (or **many-valued**). The number of function elements of $f(z)$ with the same center is at most †countably infinite, so the value of $f(z)$ at a point is a countable set (**Poincaré-Volterra theorem**). By introducing a †Riemann surface instead of the complex plane as the domain of definition of an analytic function, we can regard multiple-valued analytic functions as single-valued functions defined on a suitable Riemann surface (\rightarrow 367 Riemann Surfaces).

Let $f(z)$ be an analytic function and $P(z; a)$ be a function element belonging to $f(z)$, where a is a point of a domain D . The set of all function elements obtained from $P(z; a)$ by every possible analytic continuation along all curves in D is called a **branch** of $f(z)$ in D determined by $P(z; a)$. When D coincides with the whole complex plane, the branch of $f(z)$ in D is the function $f(z)$ itself. A function holomorphic in a domain D can be expanded in a power series with any point of D as its center, and the set of these power series (function elements) constitutes a branch of an analytic function.

If analytic continuations of a function element are possible along all curves in D , then the analytic continuations along two †homotopic curves in D lead to the same result (**monodromy theorem**). In particular, if D is †simply connected and if analytic continuations of $P(z; a)$ are possible along all curves in D starting from a , then the branch of $f(z)$ in D determined by $P(z; a)$ is single-valued.

K. Invariance Theorem of Analytic Relations

Suppose the following four conditions hold: (1) $F(z, w)$ is a holomorphic function of two variables for $z \in \Delta_1$ and $w \in \Delta_2$, where Δ_1, Δ_2 are domains in the complex plane. (2) A curve $C: z = z(s)$ ($0 \leq s \leq 1, z(0) = a, z(1) = b$) and two sets of function elements $P(z; z(s))$ and $Q(z; z(s))$ defined for every s ($0 \leq s \leq 1$) are given. (3) $P(z; a)$ and $Q(z; a)$ can be continued analytically along C using $P(z; z(s))$ and $Q(z; z(s))$, respectively. (4) There exists a positive number $R(s)$ for every s ($0 \leq s \leq 1$) such that, if $|z - z(s)| < R(s)$, the values of $P(z; z(s))$ and $Q(z; z(s))$ belong to Δ_1 and Δ_2 , respectively. Under these conditions, if $F(P(z; a), Q(z; a)) = 0$ holds for $|z - a| < R(0)$, then $F(P(z; b), Q(z; b)) = 0$ holds for $|z - b| < R(1)$. In other words, an analytic

relation between function elements belonging to two analytic functions that holds in a neighborhood of the starting point of a curve C is conserved for function elements with center at the terminal point b of C . This is called the **invariance theorem of analytic relations**. The same statement is valid for relations among more than two analytic functions and their derivatives (differential equations).

L. Inverse Functions

Suppose that $P(z; a)$ ($a \neq \infty$) belongs to an analytic function $f(z)$ and $P'(a; a) \neq 0$. We consider the inverse function of $P(z; a)$ in a neighborhood of a and let $\mathfrak{B}(w; \alpha)$ ($\alpha = P(a; a)$) be its expansion as the power series in $w - \alpha$. We call $\mathfrak{B}(w; \alpha)$ the **inverse function element** (or simply **inverse element**) of $P(z; a)$ and the analytic function determined by $\mathfrak{B}(w; \alpha)$ the **inverse analytic function** (or simply **inverse function**) of $f(z)$. The inverse function is completely determined by $f(z)$ and is independent of the choice of $P(z; a)$. For example, analytic functions represented by \sqrt{w} or $\log w$ are defined as the inverse function of z^2 or e^z , respectively.

M. Singularities of Analytic Functions

Hereafter, when we speak of a curve $C: z = z(s)$ ($0 \leq s \leq 1$), it is always supposed that C is a curve in the complex plane starting at a and ending at ω . Let K_r be the open disk $|z - \omega| < r$; we denote by C_r the connected component of $C \cap K_r$ that contains ω . If analytic continuations of $P(z; a)$ are possible along any subarc of C with a terminal point arbitrarily near ω but impossible along the whole C , we say that the analytic continuation of $P(z; a)$ along C defines a **singularity** Ω of the coordinate ω , and that Ω lies over ω . For example, if $P(z; a)$ has a finite radius of convergence, for a suitable point ω on the circle of convergence the analytic continuation of $P(z; a)$ along the radius $a\omega$ defines a singularity over ω . Now take a point z_r on C_r and denote by $F_r(z)$ the branch of an analytic function determined by $P(z; z_r)$ in K_r . Let Ω be a singularity determined by C and $P(z; a)$, and suppose that we are given another singularity Ω^* over ω determined by C^* and $P^*(z; a^*)$. If they define the same branch $F_r(z)$ for every K_r , by definition, we put $\Omega = \Omega^*$. Thus $F_r(z)$ defines an †unramified covering surface W_r of $K_r - \{\omega\}$, and it is single-valued on W_r .

Singularities are classified according to the geometric structure of W_r and the value distribution of $F_r(z)$ on it. First, if W_r has no †relative boundary over $0 < |z - \omega| < r$ for a suitable r , then Ω is called an **isolated singularity** of the

analytic function. In this case, the number k of points of W_r lying over a point z in $K_r - \{\omega\}$ is constant. If $k = \infty$, W_r has a [†]logarithmic branch point over ω , and Ω is called a **logarithmic singularity**. If $k < \infty$, $F_r(z)$ can be represented as a single-valued holomorphic function in $0 < |t| < r^{1/k}$ by putting $z = \omega + t^k$. In this case, if we introduce an additional point P_0 corresponding to $z = \omega$, then $W_r \cup \{P_0\}$ has only an [†]algebraic branch point over ω . Now, taking into account the value of $w = F_r(z)$, we call Ω an **algebraic singularity** if $\lim w$ exists. In this case, we have $F_r(z) = \sum_{n=l}^{\infty} c_n t^n$, and if we admit analytic continuations in the wider sense (→ Section O), $P(z; a)$ is analytically continuable along the whole C .

N. The Natural Boundary

Given a domain D and an analytic function $f(z)$ holomorphic in D , if all boundary points of D are singularities of $f(z)$ and $f(z)$ is not continuable to the exterior of D , the boundary of D is called the **natural boundary** of $f(z)$. This phenomenon was first discovered for [†]elliptic modular functions. Many results are known about power series for which the circle of convergence is the natural boundary (→ 339 Power Series). For any given domain D in \mathbb{C} , there exists an analytic function whose natural boundary is the boundary of D . The original proof of this fact, given by Weierstrass, contained a defect that was corrected by J. Besse.

O. Analytic Continuation in the Wider Sense

Let two [†]Laurent series (with parameter t) $z = P(t) = \sum_{n=k}^{\infty} a_n t^n$ and $w = Q(t) = \sum_{n=l}^{\infty} b_n t^n$ (k and l are integers, and $a_k b_l \neq 0$) converge in $0 < |t| < r$, and let $(P(t_1), Q(t_1)) \neq (P(t_2), Q(t_2))$ if $t_1 \neq t_2$; then we say that the pair (P, Q) defines a **function element in the wider sense**. If a change of parameter $\tau = r_1 t + r_2 t^2 + \dots$ ($r_1 \neq 0$ and the radius of convergence > 0) gives $P(t) = \Pi(\tau)$, $Q(t) = K(\tau)$, we say that (Π, K) and (P, Q) define the same function element. By a suitable choice of parameter, any function element can be given in the form $z = t^k + a$ (or $a = t^{-k}$), $w = \sum_{n=l}^{\infty} b_n t^n$, and the elimination of t gives the representation of w as a [†]Puiseux series of z . So if $k = 1$ and $l \geq 0$, it reduces to a holomorphic function element. When $k = 1$, with $l < 0$ not excluded, the above element is called a **rational element**. If $k > 1$ it is called a **ramified element**, and if $l < 0$ it is called a **polar element**.

If P', Q' are the direct analytic continuations of P and Q at t_0 ($0 < |t_0| < r$), i.e., their Taylor expansions at t_0 , the function element (P', Q') is called a direct analytic continuation of (P, Q) ,

which is also considered its own direct analytic continuation. For a fixed r , the set of all direct continuations of (P, Q) thus obtained is called an **analytic neighborhood** of (P, Q) , and these neighborhoods define a topology in the set of all function elements. A curve in this topological space is called an **analytic continuation in the wider sense**, and a [†]connected component of this space is called an **analytic function in the wider sense**. An analytic function in the wider sense is a set of function elements in the wider sense, but it can also be regarded as a function $w = f(z)$ (with an independent variable z and a dependent variable w) defined by each function element $p(z, w): z = P(t), w = Q(t)$. An analytic continuation in the wider sense,

$$p(s) = p(z, w; s);$$

$$z = z(s) + t^{k(s)}, \quad w = \sum_{n=l(s)}^{\infty} c_n(s) t^n, \quad 0 \leq s \leq 1,$$

is sometimes called an analytic continuation along the curve $C: z = z(s)$ ($0 \leq s \leq 1$) in the complex plane. If all $p(s)$ are holomorphic function elements, this coincides with the analytic continuation along C in the original sense, but if this is not the case, $p(0)$ and C do not necessarily determine $p(1)$ uniquely. Actually, an analytic function in the wider sense is just an analytic function in the original sense with at most a countable number of ramified or polar elements added.

P. Singularities of Analytic Functions in the Wider Sense

Suppose the following three conditions hold: (1) For every point on C except ω , that is, for $z(s)$ ($0 \leq s < 1$), a function element in the wider sense $p(z, w; s)$ is given. (2) For every λ (< 1), $p(z, w; s)$ ($0 \leq s \leq \lambda$) constitutes an analytic continuation in the wider sense. (3) It is impossible to find a function element $p(z, w; 1)$ such that $p(z, w; s)$ ($0 \leq s \leq 1$) is an analytic continuation in the wider sense. When these three conditions are satisfied, we say that $p(z, w; s)$ ($0 \leq s < 1$) defines a **transcendental singularity** Ω with ω as its coordinate. The method of determining a branch $w = F_r(z)$ in an open disk with center ω is completely parallel to the case of holomorphic analytic functions. Because of the appearance of function elements in the wider sense in $F_r(z)$, the covering surface W_r of K_r defined by $F_r(z)$ may have algebraic branch points. If W_r has a logarithmic branch point over ω , Ω is called a **logarithmic singularity**. If W_r has no point over ω for suitable r , Ω is called a **direct transcendental singularity**; otherwise, it is called an **indirect transcendental singularity**. The logarithmic singularities are

direct singularities. The inverse function of $z = w \sin w$ has a direct singularity over $z = \infty$ that is not logarithmic, and the inverse function of $z = (\sin w)/w$ has an indirect singularity over $z = 0$. Taking into account the value of $w = F_r(z)$, if the \dagger cluster set of F_r at $\Omega: S_\Omega = \bigcap_{r>0} \{F_r(z)\}$ consists of only one point, it is an **ordinary singularity**; if not, it is an **essential singularity**.

Q. History

A function of a complex variable is **monogenic** in the sense of A. L. Cauchy if it is differentiable at every point of its domain of definition. It was B. Riemann who succeeded in developing Cauchy's concept. Riemann considered an analytic function as a function defined on a \dagger Riemann surface, that is, a 1-dimensional complex analytic manifold. On the other hand, Weierstrass constructed the theory of analytic functions starting from power series. When we speak of single-valued functions defined in a domain of the complex plane, the monogenic functions of Cauchy and the analytic functions of Weierstrass are identical. Although the analytic functions are very special functions, the study of complex analytic functions is usually called the theory of functions of a complex variable, or simply the theory of functions.

By considering the following point set C , which is more general than a domain, E. Borel showed that a monogenic function on C is not necessarily holomorphic in the ordinary sense. Take a countable dense subset in a subdomain D' of a domain D and a double sequence of positive numbers $\{r_n^{(h)}\}$. Put $S_n^{(h)} = \{z \mid |z - z_n| < r_n^{(h)}\}$ and $C^{(h)} = D - \bigcup_{n=1}^{\infty} S_n^{(h)}$. By a suitable choice of $r_n^{(h)}$, we can suppose that the $C^{(h)}$ are connected and monotone increasing with respect to h . Put $C = \bigcup_{h=1}^{\infty} C^{(h)}$. A function defined in C is by definition **monogenic** if it is differentiable in $C^{(h)}$ for every h . For such a monogenic function, Cauchy's \dagger integral formula in a generalized form holds, and the function is infinitely differentiable. If $f(z)$ and $g(z)$ are monogenic in C and coincide on a curve in C , then they are identical in C . Let D be the set $\{z \mid 0 < \operatorname{Re} z < 1, 0 < \operatorname{Im} z < 1\}$ and $\{z_n\}$ be all rational points in D ($z_n = (p + iq)/m$). For a natural number h , we define $C^{(h)}$ to be the set D minus the union of open disks with radius $\exp(-e^{m^2})/h$ and center $(p + iq)/m$. The function

$$f(z) = \sum_{m=1}^{\infty} \sum_{p=0}^m \sum_{q=0}^m \frac{\exp(-e^{m^4})}{z - (p + iq)/m}$$

is monogenic in C in the above-mentioned sense, but not holomorphic in C . The study of

these functions developed into the theory of \dagger quasi-analytic functions.

The concept of \dagger analytic functions of several complex variables can also be defined analogously to the case of one variable. Then non-uniformizable singularities appear that lead to a generalization of the concept of \dagger manifolds (\rightarrow 23 Analytic Spaces).

References

- [1] E. Borel, *Leçons sur les fonctions monogènes uniformes d'une variable complexe*, Gauthier-Villars, 1917.
 - [2] A. Hurwitz and R. Courant, *Vorlesungen über allgemeine Funktionentheorie und elliptische Funktionen*, Springer, fourth edition, 1964.
 - [3] L. Bieberbach, *Lehrbuch der Funktionentheorie*, Teubner, I, 1921; II, 1927 (Johnson Reprint Co., 1969).
 - [4] E. C. Titchmarsh, *The theory of functions*, Oxford Univ. Press, second edition, 1939.
 - [5] C. Carathéodory, *Funktionentheorie I, II*, Birkhäuser, 1950; English translation, *Theory of functions*, Chelsea, I, 1958; II, 1960.
 - [6] S. Saks and A. Zygmund, *Analytic functions*, Warsaw, 1952.
 - [7] L. V. Ahlfors, *Complex analysis*, McGraw-Hill, 1953; third edition, 1979.
 - [8] H. Behnke and F. Sommer, *Theorie der analytischen Funktionen einer komplexen Veränderlichen*, Springer, second edition, 1962.
 - [9] H. Kneser, *Funktionentheorie*, Vandenhoeck & Ruprecht, 1958.
 - [10] E. Hille, *Analytic function theory*, Ginn, I, 1959; II, 1962.
 - [11] H. P. Cartan, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Hermann, 1961; English translation, *Elementary theory of analytic functions of one or several complex variables*, Addison-Wesley, 1963.
 - [12] A. Dinghas, *Vorlesungen über Funktionentheorie*, Springer, 1961.
 - [13] R. Nevanlinna and V. Paatero, *Einführung in die Funktionentheorie*, Birkhäuser, 1964; English translation, *Introduction to complex analysis*, Addison-Wesley, 1968.
 - [14] M. Heins, *Complex function theory*, Academic Press, 1968.
 - [15] B. A. Fuks and V. I. Levin, *Functions of a complex variable and some of their applications I, II*, Pergamon, 1961. (Original in Russian, 1951.)
 - [16] W. Rudin, *Real and complex analysis*, McGraw-Hill, 1966.
 - [17] S. Lang, *Complex analysis*, Addison-Wesley, 1977.
- As to the Looman-Men'shov theorem,

[18] D. Men'shov (Menchoff), Les conditions de monogénéité, *Actualités Sci. Ind.*, Hermann, 1936.

[19] D. Men'shov (Menchoff), Sur la généralisation des conditions de Cauchy-Riemann, *Fund. Math.*, 25 (1935), 59–97.

[20] S. Saks, *Theory of the integral*, Warsaw, 1937, 188–201.

For a new proof of Cauchy's integral theorem,

[21] E. Artin, On the theory of complex functions, *Notre Dame mathematical lectures*, 1944, 57–70. Also → [7, 17].

For problems,

[22] L. Volkovyskii, G. Lunts, and I. Aramovich, A collection of problems on complex analysis, Addison-Wesley, 1965.

For analytic continuation,

[23] H. Weyl, *Die Idee der Riemannschen Fläche*, Teubner, 1913; third edition, 1955; English translation, *The concept of a Riemann surface*, Addison-Wesley, 1964. Also → [1–17].

For the history and concepts of analytic functions,

[24] G. Julia, *Essai sur le développement de la théorie des fonctions de variables complexes*, Gauthier-Villars, 1933.

199 (IV.12)

Homogeneous Spaces

A. General Remarks

Let M be a \dagger differentiable manifold. If a \dagger Lie group G acts \dagger transitively on M as a \dagger Lie transformation group, the manifold M is said to be a **homogeneous space** having G as its transformation group (→ 431 Transformation Groups). The \dagger stabilizer (isotropy subgroup) H_x of G at a point x of M is a closed subgroup of G , and a one-to-one correspondence between G/H_x and M preserving the action of G is defined by associating the element sH_x ($s \in G$) of G/H_x with the point $s(x)$ of M . This correspondence is a \dagger diffeomorphism between the manifold M and the quotient manifold G/H_x if the number of connected components of G is at most countable. Under this condition we may therefore identify a homogeneous space M with the quotient manifold G/H of a Lie group G by a closed Lie subgroup H (→ 249 Lie Groups). However, H is not uniquely determined by M , and it may be replaced by $H_{s(x)} = sH_x s^{-1}$ ($s \in G$). Each element h of the stabilizer H_x at a point x induces a linear transformation \tilde{h} on the \dagger tangent space V_x of M at the point x . The set \tilde{H}_x of all \tilde{h} is called the **linear isotropy group** at the point x .

If we represent the homogeneous space M as G/H , we obtain the canonical map $\pi: s \rightarrow sH$ of G onto M , which we call the **projection** of G onto M . Let \mathfrak{g} be the \dagger Lie algebra of G , and \mathfrak{h} the Lie subalgebra corresponding to the closed subgroup H . When we identify \mathfrak{g} with the tangent space at the identity element e of G and \mathfrak{h} with its subspace, the projection π induces a linear isomorphism of $\mathfrak{g}/\mathfrak{h}$ with the tangent space V_x of M at the point $x = \pi(e)$. The \dagger adjoint representation of G gives rise to a linear representation $h \rightarrow \text{Ad}(h)$ modulo \mathfrak{h} of the group H on the linear space $\mathfrak{g}/\mathfrak{h}$. Through the linear isomorphism between $\mathfrak{g}/\mathfrak{h}$ and the tangent space V_x defined by the projection π , this representation of H is equivalent to the one which associates with h the linear transformation \tilde{h} defined by h on the tangent space V_x .

The homogeneous space G/H is said to be **reductive** if there exists a linear subspace \mathfrak{m} of \mathfrak{g} such that $\mathfrak{g} = \mathfrak{h} + \mathfrak{m}$ (direct sum as linear spaces) and $(\text{Ad } H)\mathfrak{m} \subset \mathfrak{m}$. H is said to be **reductive** in \mathfrak{g} if the representation $h \rightarrow \text{Ad}(h)$ of H in \mathfrak{g} is \dagger completely reducible.

If a \dagger tensor field P on the homogeneous space $M = G/H$ is G -invariant (namely, invariant under the transformations defined by the elements of G), then the value of P at the point $x = \pi(e)$ is a \dagger tensor over the tangent space V_x at x which is invariant under the linear isotropy group \tilde{H} . Conversely, such a tensor over V_x is uniquely extended to a G -invariant tensor field on M . If G/H is reductive, then G -invariant tensor fields over M are in one-to-one correspondence with \tilde{H} -invariant tensors over \mathfrak{m} . For instance, if H is compact, then H is reductive in \mathfrak{g} and an \tilde{H} -invariant positive definite quadratic form on \mathfrak{m} defines a G -invariant \dagger Riemannian metric on G/H .

We say that the homogeneous space $M = G/H$ is a **Riemannian (linearly connected, complex Hermitian, Kähler) homogeneous space** if there exists on M a G -invariant Riemannian metric (\dagger linear connection, \dagger Hermitian metric, \dagger Kähler metric). Concerning such homogeneous spaces, there are various results on their structures and geometric properties [1–5] (→ 412 Symmetric Riemannian Spaces and Real Forms; 427 Topology of Lie Groups and Homogeneous Spaces).

B. Examples

Stiefel Manifold. A k -frame ($1 \leq k \leq n$) in a real n -dimensional Euclidean vector space \mathbf{R}^n is an ordered system consisting of k linearly independent vectors. If we regard the real \dagger general linear group of degree n , $GL(n, \mathbf{R})$, as the regular linear transformation group of \mathbf{R}^n , $GL(n, \mathbf{R})$ acts transitively on the set $V'_{n,k}(\mathbf{R})$ of

all k -frames in \mathbf{R}^n . Therefore, if H denotes the subgroup of $GL(n, \mathbf{R})$ consisting of the elements which leave fixed a given k -frame v_0^k , we may identify the set $V'_{n,k}$ and the quotient set $GL(n, \mathbf{R})/H$. Transferring the differentiable manifold structure of $GL(n, \mathbf{R})/H$ to $V'_{n,k}$ through this identification, we see that $V'_{n,k}(\mathbf{R}) = GL(n, \mathbf{R})/H$ becomes a homogeneous space (the differentiable manifold structure of $V'_{n,k}$ is defined independently of the choice of v_0^k). The space $V'_{n,k}(\mathbf{R})$ is called the **(real) Stiefel manifold of k -frames in \mathbf{R}^n** .

A k -frame is called an **orthogonal k -frame** if the vectors belonging to the frame are of length 1 and are orthogonal to each other. The set $V_{n,k}(\mathbf{R})$ of all orthogonal k -frames is a submanifold of $V'_{n,k}(\mathbf{R})$. The t orthogonal group $O(n)$ acts transitively on $V'_{n,k}(\mathbf{R})$, which is a homogeneous space represented as $V_{n,k}(\mathbf{R}) = O(n)/I_k \times O(n-k)$. The manifold $V_{n,1}(\mathbf{R})$ is actually the $(n-1)$ -dimensional sphere. We call $V_{n,k}(\mathbf{R})$ the **(real) Stiefel manifold of orthogonal k -frames** (or simply **Stiefel manifold**). The **complex Stiefel manifold** $V_{n,k}(\mathbf{C}) = U(n)/I_k \times U(n-k)$ is defined analogously.

Grassmann Manifold. Let $M_{n,k}(\mathbf{R})$ ($1 \leq k \leq n$) be the set of all k -dimensional linear subspaces of \mathbf{R}^n . The group $O(n)$ acts transitively on $M_{n,k}(\mathbf{R})$, so that we may put $M_{n,k}(\mathbf{R}) = O(n)/O(k) \times O(n-k)$. Here $O(k)$ and $O(n-k)$ are identified with the subgroups of $O(n)$ consisting of all elements leaving fixed every point of a fixed $(n-k)$ -dimensional subspace and of its orthogonal complement, respectively. In this way, $M_{n,k}(\mathbf{R})$ is a homogeneous space, which we call the **(real) Grassmann manifold**. The t proper orthogonal group $SO(n)$ acts transitively on $M_{n,k}(\mathbf{R})$, and $M_{n,k}(\mathbf{R})$ may be represented as a homogeneous space having $SO(n)$ as its transformation group. It follows that $M_{n,k}(\mathbf{R})$ is connected. The homogeneous space $\tilde{M}_{n,k}(\mathbf{R}) = SO(n)/SO(k) \times SO(n-k)$ is called the **Grassmann manifold formed by oriented subspaces**. $M_{n,1}(\mathbf{R})$ and $\tilde{M}_{n,1}(\mathbf{R})$ may be identified with the $(n-1)$ -dimensional real projective space and the $(n-1)$ -dimensional sphere, respectively.

Applying the above process for real Grassmann manifolds to the complex Euclidean vector space \mathbf{C}^n instead of \mathbf{R}^n , we see that the set $M_{n,k}(\mathbf{C})$ of all k -dimensional linear subspaces in \mathbf{C}^n is a homogeneous space with the t unitary group $U(n)$ of degree n as its transformation group, and we represent it as $U(n)/U(k) \times U(n-k)$. This space is called the **complex Grassmann manifold**. The manifold $M_{n,k}(\mathbf{C})$ is a simply connected complex manifold and has a cellular decomposition as a t CW complex whose cells are t Schubert varieties (\rightarrow 56 Characteristic Classes E). On

the other hand, $M_{n,k}(\mathbf{C})$ may be regarded as the set of all $(k-1)$ -dimensional linear subspaces in the $(n-1)$ -dimensional complex projective space. Then, by using the t Plücker coordinates of these subspaces, $M_{n,k}(\mathbf{C})$ can be realized as an t algebraic variety without singularity in the projective space of dimension

$$\binom{n}{k} - 1 \quad (\rightarrow 90 \text{ Coordinates B}).$$
 Sometimes

$M_{n,k}(\mathbf{R})$ is denoted by $G_{n,k}(\mathbf{R})$ or $G(n, k)$. In the same way, the homogeneous space represented as $Sp(n)/Sp(k) \times Sp(n-k)$ is called the t quaternion Grassmann manifold and is denoted by $M_{n,k}(\mathbf{H})$.

Flag Manifold. Let k_1, \dots, k_r be a sequence of integers such that $n > k_1 > \dots > k_r > 0$, and let $F(k_1, \dots, k_r)$ be the set of all monotone sequences $V_1 \supset \dots \supset V_r$, where V_i ($i = 1, \dots, r$) is a k_i -dimensional linear subspace in \mathbf{R}^n . For the two sequences $V_1 \supset \dots \supset V_r$ and $V'_1 \supset \dots \supset V'_r$ belonging to $F(k_1, \dots, k_r)$, there exists an element $s \in GL(n, \mathbf{R})$ such that $s(V_i) = V'_i$ ($i = 1, \dots, r$). Therefore $F(k_1, \dots, k_r)$ is a homogeneous space with $GL(n, \mathbf{R})$ as its transformation group, and is called the **proper flag manifold**. Since the unitary group $U(n)$ of degree n acts transitively on it, $F(k_1, \dots, k_r)$ is also regarded as a homogeneous space admitting $U(n)$ as its transformation group. In this case, putting $F(k_1, \dots, k_r) = U(n)/H$, H is isomorphic to the direct product $U(k_1 - k_2) \times U(k_2 - k_3) \times \dots \times U(k_r)$. In particular, when $r = n - 1$, $k_i = n - i$, the homogeneous space is the quotient space of the compact Lie group $U(n)$ by a maximal t torus T . In general, the quotient space G/T of a compact connected Lie group G by a maximal torus of G is called a **flag manifold**. If G acts effectively on G/T , G is a t semisimple compact Lie group. The complex Lie group $G^{\mathbf{C}}$ is a Lie transformation group of t biregular transformations which acts transitively on the flag manifold G/T , a simply connected Kähler homogeneous space. Here $G^{\mathbf{C}}$ is a complex Lie group having G as a maximal compact subgroup. If B is a maximal t solvable Lie subgroup (t Borel subgroup) of $G^{\mathbf{C}}$, G/T is represented as $G^{\mathbf{C}}/B$.

References

- [1] A. Borel and R. Remmert, Über kompakte homogene Kählersche Mannigfaltigkeiten, Math. Ann., 145 (1961–1962), 429–439.
- [2] Y. Matsushima, Espaces homogènes de Stein des groupes de Lie complexes, Nagoya Math. J., 16 (1960), 205–218.
- [3] K. Nomizu, Invariant affine connections on homogeneous spaces, Amer. J. Math., 76 (1954), 33–65.

[4] H. C. Wang, Closed manifolds with homogeneous complex structure, *Amer. J. Math.*, 76 (1954), 1–32.

[5] N. R. Wallach, *Harmonic analysis on homogeneous spaces*, Dekker, 1973.

[6] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Academic Press, 1978.

200 (III.24) Homological Algebra

A. General Remarks

Homological algebra is a new branch of mathematics that developed rapidly after World War II. The introduction of the theory was motivated by the observation that some algebraic ideas and mechanisms that arose in the development of algebraic topology, in particular, homology theory, can provide powerful tools for treating from a unified viewpoint various problems in algebra that previously were treated differently. One of its characteristic features lies in emphasizing, from the standpoint of categories and functors (\rightarrow 52 Categories and Functors), the functional structure of the objects to be investigated rather than their inner structure. Thus the theory of derived functors constitutes the main theme of homological algebra. This new theory turned out to have wide applications in other areas of mathematics, and the philosophy embodied in the theory has been influential in the general progress of mathematics. For general references \rightarrow [2, 5, 6]; for sheaf cohomology \rightarrow [3, 4, 8].

B. Graded Modules and Graded Objects

Let A be a \dagger ring with unity element and X be a \dagger unitary A -module. If we are given a sequence of A -submodules X_n ($n \in \mathbb{Z}$) such that $X = \sum_{n \in \mathbb{Z}} X_n$ (\dagger direct sum), we call X a **graded A -module** and X_n the **component of degree n** of X . Each element x of a graded A -module X has a unique representation $x = \sum_{n \in \mathbb{Z}} x_n$ ($x_n \in X_n$); we call x_n the **component of degree n** of x . An A -submodule Y of a graded A -module X is called **homogeneous** if $x \in Y$ implies $x_n \in Y$ ($n \in \mathbb{Z}$). In this case, $Y = \sum_n Y_n$ and the quotient module $X/Y = \sum_n X_n/Y_n$ are graded A -modules, where $Y_n = Y \cap X_n$. Let $X = \sum_n X_n$ and $Y = \sum_n Y_n$ be graded A -modules and $f: X \rightarrow Y$ be an A -homomorphism. If there is a fixed integer p such that $f(X_n) \subset Y_{n+p}$ for any $n \in \mathbb{Z}$, f is called an **A -homomorphism of degree p** . In this case,

$\text{Ker } f = \sum_n \text{Ker } f_n$ and $\text{Im } f = \sum_n \text{Im } f_{n-p}$ are homogeneous A -submodules of X and Y , respectively, where $f_n: X_n \rightarrow Y_{n+p}$ is the restriction of f on X_n .

Sometimes, by a graded module we mean only a sequence $\{X_n\}$ of A -modules X_n , without considering the direct sum $\sum_n X_n$. Similarly, we have the notion of a **graded object** $\{C_n\}$ in any category \mathcal{C} .

C. Chain Complexes and Homology Modules

By a **chain complex** (X, ∂) over A we mean a graded A -module $X = \sum_n X_n$ together with an A -homomorphism $\partial: X \rightarrow X$ of degree -1 such that $\partial \circ \partial = 0$. Hence a chain complex over A is completely determined by a sequence

$$\dots \rightarrow X_{n+1} \xrightarrow{\partial_{n+1}} X_n \xrightarrow{\partial_n} X_{n-1} \rightarrow \dots$$

of A -modules and A -homomorphisms such that $\partial_n \circ \partial_{n+1} = 0$ for all n . We call ∂ the **boundary operator**. For a chain complex (X, ∂) , we write $\text{Ker } \partial = Z(X)$, $\text{Ker } \partial_n = Z_n(X)$, $\text{Im } \partial = B(X)$, $\text{Im } \partial_{n+1} = B_n(X)$. Then $Z(X) = \sum_n Z_n(X)$, $B(X) = \sum_n B_n(X)$ are homogeneous submodules of X , called the **module of cycles** and the **module of boundaries**, respectively. $B(X)$ is a homogeneous submodule of $Z(X)$, and the quotient modules $Z(X)/B(X)$, $Z_n(X)/B_n(X)$ are denoted by $H(X)$, $H_n(X)$, respectively. We call $H(X) = \sum_n H_n(X)$ the **homology module** of the chain complex (X, ∂) .

If (X, ∂) , (Y, ∂') are chain complexes over A , an A -homomorphism $f: X \rightarrow Y$ of degree 0 satisfying $\partial' \circ f = f \circ \partial$ (i.e., $\partial'_n f_n = f_{n-1} \partial_n$) is called a **chain mapping** of X to Y . For a chain mapping f , we have $f(Z_n(X)) \subset Z_n(Y)$, $f(B_n(X)) \subset B_n(Y)$, and hence f induces an A -homomorphism $f_*: H(X) \rightarrow H(Y)$ of degree 0, which is called the **homological mapping** induced by f . We have $(1_X)_* = 1_{H(X)}$, and $(g \circ f)_* = g_* \circ f_*$ for chain mappings $f: X \rightarrow Y$ and $g: Y \rightarrow Z$.

Let $f, g: X \rightarrow Y$ be two chain mappings. If there is an A -homomorphism $D: X \rightarrow Y$ of degree $+1$ such that $f - g = D \circ \partial + \partial' \circ D$, we say that f is **chain homotopic** to g and write $f \simeq g$; D is called a **chain homotopy** of f to g . If f is chain homotopic to g , we have $f_* = g_*: H(X) \rightarrow H(Y)$. For chain complexes X and Y , if there are chain mappings $f: X \rightarrow Y$ and $g: Y \rightarrow X$ such that $f \circ g \simeq 1_Y$ and $g \circ f \simeq 1_X$, we say that X is **chain equivalent** to Y . In this case $f_*: H(X) \rightarrow H(Y)$ is an isomorphism and $g_*: H(Y) \rightarrow H(X)$ is its inverse.

Let (X, ∂) be a chain complex over A and $Y = \sum_n Y_n$ be a homogeneous A -submodule of X such that $\partial Y \subset Y$. Then Y and X/Y are chain complexes over A with the boundary operators induced by ∂ . Y is called a **chain**

subcomplex of X , and X/Y is called the **quotient chain complex** of X by Y or the **relative chain complex** of $X \bmod Y$. For a chain complex X and its subcomplex Y we have an exact sequence $0 \rightarrow Y \xrightarrow{i} X \xrightarrow{j} X/Y \rightarrow 0$, where i is the \dagger canonical injection and j the \dagger canonical surjection.

Let (W, ∂') , (X, ∂) , (Y, ∂'') be chain complexes over A , and $f: W \rightarrow X$, $g: X \rightarrow Y$ be chain mappings such that $0 \rightarrow W \xrightarrow{f} X \xrightarrow{g} Y \rightarrow 0$ is exact. Then an A -homomorphism $\partial_*: H(Y) \rightarrow H(W)$ of degree -1 , called the **connecting homomorphism**, is defined by $\partial_*(y + B(Y)) = f^{-1} \circ \partial \circ g^{-1}(y) + B(W)$ ($y \in Z(Y)$), and we have the **exact sequence of homology**:

$$\begin{aligned} \dots \xrightarrow{\partial_*} H_n(W) \xrightarrow{f_*} H_n(X) \xrightarrow{g_*} H_n(Y) \\ \xrightarrow{\partial_*} H_{n-1}(W) \xrightarrow{f_*} H_{n-1}(X) \xrightarrow{g_*} H_{n-1}(Y) \xrightarrow{\partial_*} \dots \end{aligned}$$

For a \dagger commutative diagram

$$\begin{array}{ccccc} 0 \rightarrow W & \rightarrow & X & \rightarrow & Y \rightarrow 0 \\ & \downarrow \varphi & \downarrow & \downarrow \psi & \\ 0 \rightarrow W' & \rightarrow & X' & \rightarrow & Y' \rightarrow 0 \end{array} \quad (1)$$

consisting of chain complexes and chain mappings in which each row is exact, we have $\partial_* \circ \psi_* = \varphi_* \circ \partial_*: H(Y) \rightarrow H(W')$.

For the \dagger inductive limit $\varinjlim X_\lambda$ of chain complexes X_λ over A , we have

$$H(\varinjlim X_\lambda) = \varinjlim H(X_\lambda).$$

A chain complex X is said to be **positive** if $X_n = 0$ for all $n < 0$. If X is a positive chain complex over A and M is an A -module, then we mean by an **augmentation of X over M** an A -homomorphism $\varepsilon: X_0 \rightarrow M$ such that the composition $X_1 \xrightarrow{\partial_1} X_0 \xrightarrow{\varepsilon} M$ is trivial: $\varepsilon \circ \partial_1 = 0$. A positive chain complex X together with an augmentation ε of X over M is called an **augmented chain complex** over M . It is said to be **acyclic** if the sequence

$$\dots \rightarrow X_n \xrightarrow{\partial_n} X_{n-1} \rightarrow \dots \rightarrow X_1 \xrightarrow{\partial_1} X_0 \xrightarrow{\varepsilon} M \rightarrow 0$$

is exact, namely, if $H_n(X) = 0$ ($n \neq 0$) and ε induces an A -isomorphism $H_0(X) \cong M$. In this case X is also called a **left resolution** of M .

Moreover, if each X_n is a \dagger projective A -module, X is called a **left projective resolution**. For any A -module M , there exists a left projective resolution of M .

Let $\alpha: M \rightarrow M'$ be an A -homomorphism of A -modules, and X, X' be augmented chain complexes over M, M' having augmentations $\varepsilon, \varepsilon'$, respectively. Then a chain mapping $f: X \rightarrow X'$ satisfying $\varepsilon' \circ f_0 = \alpha \circ \varepsilon$ is called a **chain mapping over α** . If X, X' are left projective resolutions of M, M' , respectively, then there exist chain mappings of X to X' over α , and any two such mappings are chain homotopic. In particular, a left projective resolution of an A -

module M is uniquely determined up to chain homotopy.

D. Tor

Given a right A -module M and a left A -module N , \mathbf{Z} -modules $\mathbf{Tor}_n^A(M, N)$ ($n = 0, 1, 2, \dots$), called the **torsion products** (or **Tor groups**), are defined as follows: Let

$$Y: \dots \rightarrow Y_n \xrightarrow{\partial_n} Y_{n-1} \rightarrow \dots \xrightarrow{\partial_1} Y_0 \xrightarrow{\varepsilon} N \rightarrow 0$$

be a projective resolution of N , and consider the chain complex

$$\begin{aligned} (M \otimes_A Y, 1 \otimes \partial): \dots \\ \rightarrow M \otimes_A Y_n \xrightarrow{1 \otimes \partial_n} M \otimes_A Y_{n-1} \rightarrow \dots \\ \xrightarrow{1 \otimes \partial_1} M \otimes_A Y_0 \rightarrow 0 \end{aligned}$$

obtained by forming the \dagger tensor product of M and Y . Then we see that the homology module $H_n(M \otimes_A Y)$ is uniquely determined for any choice of left projective resolution of N . We define $\mathbf{Tor}_n^A(M, N) = H_n(M \otimes_A Y)$. In particular, we have $\mathbf{Tor}_0^A(M, N) \cong M \otimes_A N$.

Properties of Tor. (1) If M is a \dagger flat A -module, we have $\mathbf{Tor}_n^A(M, N) = 0$ ($n = 1, 2, \dots$).

(2) An A -homomorphism $f: M_1 \rightarrow M_2$ induces a homomorphism $f_*: \mathbf{Tor}_n^A(M_1, N) \rightarrow \mathbf{Tor}_n^A(M_2, N)$. We have $(1_M)_* = 1$, and $(g \circ f)_* = g_* \circ f_*$ for $f: M_1 \rightarrow M_2, g: M_2 \rightarrow M_3$.

(3) For an exact sequence $0 \rightarrow M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3 \rightarrow 0$, we have the following **exact sequence of Tor**:

$$\begin{aligned} \dots \xrightarrow{\partial_*} \mathbf{Tor}_n^A(M_1, N) \xrightarrow{f_*} \mathbf{Tor}_n^A(M_2, N) \\ \xrightarrow{g_*} \mathbf{Tor}_n^A(M_3, N) \xrightarrow{\partial_*} \mathbf{Tor}_{n-1}^A(M_1, N) \rightarrow \dots \\ \rightarrow \mathbf{Tor}_1^A(M_3, N) \xrightarrow{\partial_*} M_1 \otimes_A N \rightarrow M_2 \otimes_A N \\ \rightarrow M_3 \otimes_A N \rightarrow 0, \end{aligned}$$

where ∂_* are the connecting homomorphisms.

(4) For a commutative diagram

$$\begin{array}{ccccc} 0 \rightarrow M_1 & \rightarrow & M_2 & \rightarrow & M_3 \rightarrow 0 \\ & \downarrow \varphi & \downarrow & \downarrow \psi & \\ 0 \rightarrow M'_1 & \rightarrow & M'_2 & \rightarrow & M'_3 \rightarrow 0 \end{array}$$

of A -modules and A -homomorphisms with exact rows, we have $\partial_* \circ \psi_* = \varphi_* \circ \partial_*$.

$$(5) \mathbf{Tor}_n^A(\sum_\lambda M_\lambda, N) \cong \sum_\lambda \mathbf{Tor}_n^A(M_\lambda, N)$$

$$(6) \mathbf{Tor}_n^A(\varinjlim M_\lambda, N) \cong \varinjlim \mathbf{Tor}_n^A(M_\lambda, N).$$

On the other hand, take a left projective resolution X of M and consider the chain complex $X \otimes_A N$. Then we have $H_n(X \otimes_A N) \cong \mathbf{Tor}_n^A(M, N)$ for $n = 0, 1, \dots$. Therefore properties similar to (1)–(6) hold with respect to the second variable N of $\mathbf{Tor}_n^A(M, N)$.

(7) If A° is a ring \dagger anti-isomorphic to A , then $\mathbf{Tor}_n^A(M, N) \cong \mathbf{Tor}_n^{A^\circ}(N, M)$. In particular,

if A is commutative, then $\text{Tor}_n^A(M, N)$ is an A -module and we have $\text{Tor}_n^A(M, N) \cong \text{Tor}_n^A(N, M)$.

(8) Let A be a \dagger principal ideal ring. Then $\text{Tor}_n^A(M, N) = 0$ ($n = 2, 3, \dots$) and $\text{Tor}_1^A(M, N)$ is also denoted by $M * {}_AN$. For an exact sequence $0 \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow 0$, we have the exact sequence $0 \rightarrow M_1 * {}_AN \rightarrow M_2 * {}_AN \rightarrow M_3 * {}_AN \rightarrow 0$. In particular, $\mathbf{Z} * {}_Z N = 0$ and $(\mathbf{Z}/n\mathbf{Z}) * {}_Z N \cong {}_nN$ ($= \{x \in N \mid nx = 0\}$).

Universal Coefficient Theorem for Homology. If (X, ∂) is a chain complex over A and N is a left A -module, then $(X \otimes {}_AN, \partial \otimes 1)$ is a chain complex. If A is a principal ideal ring and each X_n is a \dagger torsion-free A -module, then we have a formula

$$H_n(X \otimes {}_AN) \\ \cong H_n(X) \otimes {}_AN + H_{n-1}(X) * {}_AN,$$

called the **universal coefficient theorem**.

E. Double Chain Complexes

By a **double chain complex** $(X_{p,q}, \partial', \partial'')$ over A we mean a family of left A -modules $X_{p,q}$ ($p, q \in \mathbf{Z}$) together with A -homomorphisms $\partial'_{p,q}: X_{p,q} \rightarrow X_{p-1,q}$ and $\partial''_{p,q}: X_{p,q} \rightarrow X_{p,q-1}$ such that $\partial'_{p-1,q} \circ \partial'_{p,q} = \partial''_{p,q-1} \circ \partial'_{p,q} = \partial'_{p,q-1} \circ \partial''_{p,q} + \partial''_{p-1,q} \circ \partial'_{p,q} = 0$. We define the associated chain complex (X_n, ∂) by setting $X_n = \sum_{p+q=n} X_{p,q}$, $\partial_n = \sum_{p+q=n} \partial'_{p,q} + \partial''_{p,q}$. We call ∂ the **total boundary operator**, and ∂', ∂'' the **partial boundary operators**.

Given a chain complex X consisting of right A -modules and a chain complex Y consisting of left A -modules, a double chain complex $(Z_{p,q}, \partial', \partial'')$ is defined by setting $Z_{p,q} = X_p \otimes {}_AY_q$, $\partial'_{p,q} = \partial_p \otimes 1$, $\partial''_{p,q} = (-1)^p 1 \otimes \partial_q$, where ∂_p, ∂_q are the boundary operators of X, Y , respectively. It is called the **product double chain complex** of X and Y , and the homology module of its associated chain complex is denoted by $H(X \otimes {}_AY)$. With respect to this homology module, the following facts hold. If X is a left projective resolution of a right A -module M and Y is that of a left A -module N , then $H_n(X \otimes {}_AY) = \text{Tor}_n^A(M, N)$. If A is a principal ideal ring and each X_n is a torsion-free A -module, then we have the formula

$$H_n(X \otimes {}_AY) \cong \sum_{p+q=n} H_p(X) \otimes {}_AH_q(Y) \\ + \sum_{p+q=n-1} H_p(X) * {}_AH_q(Y),$$

the **Künneth theorem**.

F. Cochain Complexes

By a **cochain complex** (X, d) over A we mean a graded A -module X together with an A -

homomorphism $d: X \rightarrow X$ of degree $+1$ such that $d \circ d = 0$; d is called the **coboundary operator** or **differential**. For a cochain complex (X, d) , we denote by X^n the component of degree n of X , and by $d^n, X^n \rightarrow X^{n+1}$ the restriction of d on X^n . Then a chain complex (Y, ∂) is defined by $Y_n = X^{-n}$ and $\partial_n: Y_n \rightarrow Y_{n-1}$ is equal to $d^{-n}: X^{-n} \rightarrow X^{-n+1}$.

For a cochain complex (X, d) , we write $\text{Ker } d^n = Z^n(X)$, $\text{Ker } d = Z(X)$ ($Z(X) = \sum Z^n(X)$), $\text{Im } d^{n-1} = B^n(X)$, $\text{Im } d = B(X)$ ($B(X) = \sum B^n(X)$), and $Z^n(X)/B^n(X) = H^n(X)$, $Z(X)/B(X) = H(X)$ ($H(X) = \sum H^n(X)$). These modules $Z(X)$ ($Z^n(X)$), $B(X)$ ($B^n(X)$), and $H(X)$ ($H^n(X)$) are called the **module of cocycles**, the **module of coboundaries**, and the **cohomology module** of X , respectively. If we consider the associated chain complex (Y, ∂) of (X, d) , then $H_{-n}(Y)$ corresponds to $H^n(X)$. In this way, results on chain complexes give results on cochain complexes. Thus the concepts of **cochain mapping**, **cochain homotopy**, **cochain equivalence**, **cochain subcomplex**, and **relative cochain complex** can be defined as in the case of chain complexes in B , and we have corresponding results. In particular, given an exact sequence $0 \rightarrow W \xrightarrow{f} X \xrightarrow{g} Y \rightarrow 0$ of cochain complexes and cochain mappings, the **connecting homomorphism** $d_*: H^n(Y) \rightarrow H^{n+1}(W)$ is defined, and the **exact sequence of cohomology**

$$\dots \xrightarrow{d_*} H^n(W) \xrightarrow{f_*} H^n(X) \xrightarrow{g_*} H^n(Y) \\ \xrightarrow{d_*} H^{n+1}(W) \xrightarrow{f_*} H^{n+1}(X) \xrightarrow{g_*} H^{n+1}(Y) \xrightarrow{d_*} \dots$$

exists. For a commutative diagram

$$0 \rightarrow W \rightarrow X \rightarrow Y \rightarrow 0 \\ \downarrow \varphi \quad \downarrow \quad \downarrow \psi \\ 0 \rightarrow W' \rightarrow X' \rightarrow Y' \rightarrow 0$$

of cochain complexes and cochain mappings with exact rows, we have $d_* \circ \psi_* = \varphi_* \circ d_*$.

A cochain complex X is said to be **positive** if $X^n = 0$ for $n < 0$. If X is a positive cochain complex over A and M is an A -module, we mean by an **augmentation** of X over M an A -homomorphism $\varepsilon: M \rightarrow X^0$ such that the composition $M \xrightarrow{\varepsilon} X^0 \xrightarrow{d^0} X^1$ is trivial. If the sequence $0 \rightarrow M \xrightarrow{\varepsilon} X^0 \xrightarrow{d^0} \dots \rightarrow X^n \xrightarrow{d^n} X^{n+1} \rightarrow \dots$

is exact, X is called a **right resolution** of M . Moreover, if each X^n is an \dagger injective A -module, X is called a **right injective resolution** of M . For any A -module M , there exists a right injective resolution of M , and any two such resolutions are cochain homotopic.

G. Ext

Given left A -modules M and N , \mathbf{Z} -modules $\text{Ext}_A^n(M, N)$ ($n = 0, 1, 2, \dots$), called the **Ext**

groups, are defined as follows: Let $X: \dots \rightarrow X_n \rightarrow X_{n-1} \rightarrow \dots \rightarrow X_0 \rightarrow M \rightarrow 0$ be a projective resolution of M , and consider the cochain complex $\text{Hom}_A(X, N)$:

$$0 \rightarrow \text{Hom}_A(X_0, N) \rightarrow \dots \rightarrow \text{Hom}_A(X_{n-1}, N) \rightarrow \dots \rightarrow \text{Hom}_A(X_n, N) \rightarrow \dots$$

obtained by forming the \dagger module of A -homomorphisms. Then we can show that the cohomology module $H^n(\text{Hom}_A(X, N))$ is uniquely determined for any choice of projective resolution of M . We define $\text{Ext}_A^n(M, N) = H^n(\text{Hom}_A(X, N))$. This can also be defined as the cohomology module $H^n(\text{Hom}_A(M, Y))$ of the cochain complex $\text{Hom}_A(M, Y): 0 \rightarrow \text{Hom}_A(M, Y^0) \rightarrow \dots \rightarrow \text{Hom}_A(M, Y^{n-1}) \rightarrow \text{Hom}_A(M, Y^n) \rightarrow \dots$, where $Y: 0 \rightarrow N \rightarrow Y^0 \rightarrow \dots \rightarrow Y^{n-1} \rightarrow Y^n \rightarrow \dots$ is a right injective resolution of N . Furthermore, for a left projective resolution X of M and a right injective resolution Y of N , we see that $\text{Ext}_A^n(M, N)$ is isomorphic to the cohomology module $H^n(\text{Hom}_A(X, Y))$ of the associated cochain complex of the double cochain complex $\text{Hom}_A(X, Y) = (\text{Hom}_A(X_p, Y^q), d', d'')$, where $d'_{p,q}: \text{Hom}_A(X_p, Y^q) \rightarrow \text{Hom}_A(X_{p+1}, Y^q)$ and $d''_{p,q}: \text{Hom}_A(X_p, Y^q) \rightarrow \text{Hom}_A(X_p, Y^{q+1})$ are given by $d'_{p,q}(u) = u \circ \partial_{p+1}$, $d''_{p,q}(u) = (-1)^{p+q+1} d^q \circ u$ ($u \in \text{Hom}_A(X_p, Y^q)$) by using the boundary operator ∂ of X and the co-boundary operator d of Y .

Properties of Ext. (1) We have $\text{Ext}_A^0(M, N) \cong \text{Hom}_A(M, N)$.

(2) If M is a projective A -module or N is an injective A -module, then $\text{Ext}_A^n(M, N) = 0$ ($n = 1, 2, \dots$).

(3) An A -homomorphism $f: M_1 \rightarrow M_2$ (resp. $f: N_1 \rightarrow N_2$) induces a homomorphism $f_*: \text{Ext}_A^n(M_2, N) \rightarrow \text{Ext}_A^n(M_1, N)$ (resp. $f_*: \text{Ext}_A^n(M, N_1) \rightarrow \text{Ext}_A^n(M, N_2)$). We have $1_M^* = 1$ and $(g \circ f)^* = f^* \circ g^*$ for $f: M_1 \rightarrow M_2$, $g: M_2 \rightarrow M_3$ (resp. $1_N^* = 1$ and $(g \circ f)_* = g_* \circ f_*$ for $f: N_1 \rightarrow N_2$, $g: N_2 \rightarrow N_3$).

(4) For an exact sequence $0 \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow 0$ (resp. $0 \rightarrow N_1 \rightarrow N_2 \rightarrow N_3 \rightarrow 0$), we have the **exact sequence of Ext**:

$$\begin{aligned} 0 \rightarrow \text{Hom}_A(M_3, N) \rightarrow \text{Hom}_A(M_2, N) \\ \rightarrow \text{Hom}_A(M_1, N) \rightarrow \text{Ext}_A^1(M_3, N) \\ \rightarrow \text{Ext}_A^1(M_2, N) \rightarrow \dots \\ (\text{resp. } 0 \rightarrow \text{Hom}_A(M, N_1) \rightarrow \text{Hom}(M, N_2) \\ \rightarrow \text{Hom}(M, N_3) \rightarrow \text{Ext}_A^1(M, N_1) \\ \rightarrow \text{Ext}_A^1(M, N_2) \rightarrow \dots) \end{aligned}$$

(5)

$$\text{Ext}_A^n\left(\sum_{\alpha} M_{\alpha}, \prod_{\beta} N_{\beta}\right) = \prod_{\alpha, \beta} \text{Ext}_A^n(M_{\alpha}, N_{\beta}).$$

(6) If A is a principal ideal ring, then $\text{Ext}_A^n(M, N) = 0$ ($n = 2, 3, \dots$), and $\text{Ext}_A^1(M, N)$ is also denoted by $\text{Ext}_A(M, N)$. In particular, $\text{Ext}_{\mathbb{Z}}(\mathbb{Z}, N) = 0$, $\text{Ext}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, N) \cong N/nN$, $\text{Ext}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z}) = 0$, $\text{Ext}_{\mathbb{Z}}(M, \mathbb{Z}/n\mathbb{Z}) = \hat{M}/n\hat{M}$, where $\hat{M} = \text{Hom}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z})$.

Universal Coefficient Theorem for Cohomology

If X is a chain complex over a principal ideal ring A such that each X_n is a free A -module, then for any A -module N we have the formula

$$\begin{aligned} H^n(\text{Hom}_A(X, N)) \\ \cong \text{Hom}_A(H_n(X), N) + \text{Ext}_A(H_{n-1}(X), N), \end{aligned}$$

the **universal coefficient theorem**. This is generalized as follows: Let X be a chain complex and Y a cochain complex, both over a principal ideal ring A . Assume that each X_n is a free A -module or that each Y^n is an injective A -module. Then we have the formula

$$\begin{aligned} H^n(\text{Hom}_A(X, Y)) \\ \cong \sum_{p+q=n} \text{Hom}_A(H_p(X), H^q(Y)) \\ + \sum_{p+q=n-1} \text{Ext}_A(H_p(X), H^q(Y)) \end{aligned}$$

(\rightarrow 201 Homology Theory).

H. Complexes in Abelian Categories

We mainly consider general \dagger Abelian categories \mathcal{C} . Consideration may, however, be restricted to the \dagger category (Ab) of Abelian groups (whose \dagger objects are Abelian groups and whose \dagger morphisms are homomorphisms) or the \dagger category ${}_R\mathcal{M}$ of R -modules.

A **(cochain) complex** C in an Abelian category \mathcal{C} is a graded object $\{C^n\}$ in \mathcal{C} with **differentials** $d^n: C^n \rightarrow C^{n+1}$ subject to the condition that $d^{n+1} \circ d^n = 0$ ($n \in \mathbb{Z}$). The n th **cohomology** $H^n(C)$ of C is defined by the \dagger exact sequence $0 \rightarrow B^n(C) \rightarrow Z^n(C) \rightarrow H^n(C) \rightarrow 0$, where $B^n(C)$ and $Z^n(C)$ are objects representing $\text{Im } d^{n-1}$ and $\text{Ker } d^n$, respectively. The complex C is called **positive (negative)** if $C^n = 0$ for $n < 0$ ($n > 0$). We sometimes interchange positive superscripts and negative subscripts and write C_{-n} instead of C^n . Then the differentials become $d_n: C_n \rightarrow C_{n-1}$, and C is then called a **chain complex**. The quotient of $\text{Ker } d_n = Z_n$ by $\text{Im } d_{n+1} = B_n$ is called the n th **homology** $H_n(C)$. Negative complexes are usually described in this manner. When C^n , Z^n , B^n , and H^n are sets, as in the category ${}_R\mathcal{M}$ of R -modules, their elements are called **cochains**, **cocycles**, **coboundaries**, and **cohomology classes**, respectively. Similarly, in the group C_n of **chains**, residue classes of **cycles** ($\in Z_n$) modulo **boundaries** ($\in B_n$) are called **homology classes** ($\in H_n$).

A **morphism** (or **chain transformation**) $f: C \rightarrow C'$ is a \dagger natural transformation of the complexes considered as \dagger functors $\mathbf{Z} \rightarrow \mathcal{C}$; i.e., f is a family of morphisms $f^n: C^n \rightarrow C'^n$ ($n \in \mathbf{Z}$) satisfying $f^{n+1} \circ d^n = d'^n \circ f^n$. It induces a morphism of cohomology $H^n(C) \rightarrow H^n(C')$. A **subcomplex** of C is an equivalence class of \dagger monomorphisms $D \rightarrow C$, usually denoted by any representative D of the class. A **(chain) homotopy** between two chain transformations $f, g: C \rightarrow C'$ is a family of morphisms $h^n: C^n \rightarrow C'^{n-1}$ ($n \in \mathbf{Z}$) satisfying $f^n - g^n = h^{n+1} \circ d^n + d'^{n-1} \circ h^n$. If there exists a homotopy between f and g , then f and g induce the same morphism of cohomology. A morphism $f: C \rightarrow C'$ is called a **(chain) equivalence** if there exists a morphism $f': C' \rightarrow C$ such that $f' \circ f$ and $f \circ f'$ are homotopic to the identities of C and C' , respectively. In this case we have $H^n(C) \cong H^n(C')$. An exact sequence of complexes $0 \rightarrow C' \rightarrow C \rightarrow C'' \rightarrow 0$ gives rise to the **connecting morphisms** $H^n(C'') \rightarrow H^{n+1}(C')$ ($n \in \mathbf{Z}$), and the resulting sequence $\dots \rightarrow H^{n-1}(C'') \rightarrow H^n(C') \rightarrow H^n(C) \rightarrow H^n(C'') \rightarrow H^{n+1}(C') \rightarrow \dots$ is exact (the **exact cohomology sequence**), and similarly for homology instead of cohomology. An object $A \in \mathcal{C}$ defines a complex (also denoted by A) such that $A^0 = A$, $d^0 = 0$. A positive complex C together with a morphism $\varepsilon: A \rightarrow C$ is called a **complex over A** , and ε is the **augmentation**. A complex C over A is **acyclic** if $0 \rightarrow A \xrightarrow{\varepsilon} C^0 \rightarrow C^1 \rightarrow \dots$ is exact. An acyclic positive complex over A is called a **right resolution** of A . Let $\{C, \varepsilon\}, \{C', \varepsilon'\}$ be complexes over A and A' , respectively, and α a morphism $A \rightarrow A'$. Then a morphism $f: C \rightarrow C'$ satisfying $f \circ \varepsilon = \varepsilon' \circ \alpha$ is called a morphism over α . For a negative complex C , we define similarly augmentations $\varepsilon: C \rightarrow A$, acyclicity, left resolutions, etc.

A **bicomplex** (or **double complex**) C in \mathcal{C} consists of objects $C^{p,q}$ ($p, q \in \mathbf{Z}$) and two differentials $d_1: C^{p,q} \rightarrow C^{p+1,q}$, $d_2: C^{p,q} \rightarrow C^{p,q+1}$ subject to $d_1^2 = d_2^2 = 0$ and $d_1 d_2 = d_2 d_1$ (sometimes replaced by anticommutativity, $d_1 d_2 + d_2 d_1 = 0$). Morphisms of bicomplexes are defined as for single complexes. A bicomplex C becomes a (single) complex if we put $C^n = \sum_{p+q=n} C^{p,q}$ (when the sum exists) and define the differential d to be $d_1 + (-1)^p d_2$ on $C^{p,q}$. Then d is called the **total differential** and d_1, d_2 the **partial differentials**. On the other hand, $C_1^q = \{C^{p,q} (p \in \mathbf{Z}), d_1\}$ constitutes a complex for each q , whose cohomology $H^p(C_1^q)$ is denoted by $H_1^p(C^q)$. Then d_2 induces morphisms $H_1^p(C^q) \rightarrow H_1^p(C^{q+1})$, so that we obtain a complex $H_1^p(C)$. The cohomology of $H_1^p(C)$ is denoted by $H_2^q(H_1^p(C))$. We define $H_1^p(H_2^q(C))$ similarly. The cohomology of C with respect to the total differential is denoted simply by $H^n(C)$. Similar constructions are applied to double chain

complexes $\{C_{p,q}\}$ and further to **multiple complexes**, as we shall show in the case of bicomplexes.

Let T be a \dagger bifunctor $\mathcal{C}_1 \times \mathcal{C}_2 \rightarrow \mathcal{C}'$ and C_i be complexes in \mathcal{C}_i ($i = 1, 2$). Then $T(C_1, C_2)$ is a bicomplex in \mathcal{C}' . For instance, $\text{Hom}(C', C)$ is a positive (bipositive) complex if $C(C')$ is a positive (negative) complex in \mathcal{C} . If C, C' are complexes in $\mathcal{M}_R, {}_R\mathcal{M}$, respectively, the \dagger tensor product $C \otimes_R C'$ is a complex in (Ab) (the **product complex**). There is a canonical morphism $H_p(C) \otimes H_q(C') \rightarrow H_{p+q}(C \otimes C')$. If C_n and B_n are \dagger flat for all $n \in \mathbf{Z}$, we have the following exact sequence (**Künneth's formula**):

$$0 \rightarrow \sum_{p+q=n} H_p(C) \otimes H_q(C') \rightarrow H_n(C \otimes C') \\ \rightarrow \sum_{p+q=n-1} \text{Tor}_1(H_p(C), H_q(C')) \rightarrow 0$$

(for the definition of Tor → Section D). For $C' = A \in {}_R\mathcal{M}$, Künneth's formula reduces to the exact sequence $0 \rightarrow H_n(C) \otimes A \rightarrow H_n(C \otimes A) \rightarrow \text{Tor}_1(H_{n-1}(C), A) \rightarrow 0$ (**universal coefficient theorem**). The corresponding exact sequence for cohomology is

$$0 \rightarrow \text{Ext}^1(H_{n-1}(C), A) \rightarrow H^n(C, A) \\ \rightarrow \text{Hom}(H_n(C), A) \rightarrow 0$$

(→ Section G; 201 Homology Theory).

I. Satellites and Derived Functors

Let \mathcal{C} and \mathcal{C}' be Abelian categories. All functors in this section are \dagger additive. A \dagger covariant functor $T: \mathcal{C} \rightarrow \mathcal{C}'$ is called **exact** if T maps every exact sequence in \mathcal{C} to an exact sequence in \mathcal{C}' . T is called **half-exact**, **left exact**, or **right exact** if for every **short exact sequence** $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$, the sequence $T(A) \rightarrow T(B) \rightarrow T(C)$, $0 \rightarrow T(A) \rightarrow T(B) \rightarrow T(C)$, or $T(A) \rightarrow T(B) \rightarrow T(C) \rightarrow 0$, respectively, is exact. Similar definitions apply for \dagger contravariant functors. The functor $\text{Hom}: \mathcal{C} \times \mathcal{C} \rightarrow (\text{Ab})$ (which defines the category \mathcal{C}) is left exact in both factors. An object P is **projective** if $h_p(\cdot) = \text{Hom}(P, \cdot)$ is exact, while Q is **injective** if $h^q(\cdot) = \text{Hom}(\cdot, Q)$ is exact. If every object A admits a \dagger epimorphism from a projective object $P \rightarrow A$ (resp. \dagger monomorphism into an injective object $A \rightarrow Q$), \mathcal{C} is said to have enough projectives (injectives). An object G is called a **generator (cogenerator)** if the natural mapping $\text{Hom}(A, B) \rightarrow \text{Hom}(h_G(A), h_G(B))$ ($\text{Hom}(h^G(B), h^G(A))$) is one-to-one.

An Abelian category \mathcal{C} is called a **Grothendieck category** if (1) \mathcal{C} has a generator, (2) \dagger direct sums always exist, and (3) the identity $(\bigcup A_i) \cap B = \bigcup (A_i \cap B)$ holds for any object A , its \dagger subobject B , and a \dagger totally ordered family $\{A_i\}$ of subobjects. A Grothendieck cate-

gory has enough injectives (R. Baer, 1940, for (Ab); A. Grothendieck, 1957, for general \mathcal{C}). A monomorphism into an injective object $f: A \rightarrow Q$ is called an **injective envelope** if $\text{Im } f \cap \text{Im } g \neq 0$ for any nonzero monomorphism $g: B \rightarrow Q$. Every object A in a Grothendieck category admits an injective envelope, which is unique up to isomorphism (B. Eckmann and A. Schopf, 1953, for ${}_R\mathcal{M}$; B. Mitchell, 1960, for general \mathcal{C}).

We say that a covariant ∂ -**functor** $\mathcal{C} \rightarrow \mathcal{C}'$ is given if we have a sequence of covariant functors $T = \{T^i: \mathcal{C} \rightarrow \mathcal{C}'\}$ and the **connecting morphisms** $\partial: T^i(A') \rightarrow T^{i+1}(A')$ for an arbitrary short exact sequence $0 \rightarrow A' \rightarrow A \rightarrow A'' \rightarrow 0$ satisfying the following conditions: (i) $\partial \circ T^i(f'') = T^{i+1}(f') \circ \partial$ for a morphism f of short exact sequences; and (ii) the sequence $\dots \rightarrow T^{i-1}(A') \xrightarrow{\partial} T^i(A') \rightarrow T^i(A) \rightarrow T^i(A'') \xrightarrow{\partial} T^{i+1}(A') \rightarrow \dots$ constitutes a complex. $T = \{T^i\}$ is called a covariant ∂^* -**functor** if instead of ∂ there are given $\partial^*: T^i(A'') \rightarrow T^{i-1}(A')$ satisfying similar conditions (i*) and (ii*). By taking duals, we define the notion of contravariant ∂ - and ∂^* -functors. They are also called **connected sequences of functors**. A ∂ -(∂^*)-functor defined for $-\infty < i < +\infty$ is called a **cohomological functor (homological functor)** if the sequence in condition (ii) (resp. (ii*)) is always exact. A morphism of ∂ -functors $f: S \rightarrow T$ consists of natural transformations $f^i: S^i \rightarrow T^i$ that commute with the connecting morphisms. A ∂ -functor S defined for $a \leq i < b$ is called **universal** if for any ∂ -functor T defined in the same interval and any natural transformation $\varphi: S^a \rightarrow T^a$, there exists one and only one morphism $f: S \rightarrow T$ such that $f^a = \varphi$. Let $F: \mathcal{C} \rightarrow \mathcal{C}'$ be a covariant functor and b any positive integer. A universal covariant ∂ -functor S defined for $0 \leq i < b$ is called a **right satellite** of F if $S^0 = F$ (S is then denoted by $\{S^i F\}$). If such an S exists, then it is unique and satisfies $S^{i+1}(F) = S^1(S^i F)$. If \mathcal{C} has enough injective objects, the right satellites always exist, and if F is left exact, then $\{S^i F\}$ is a cohomological functor. The universality of ∂^* -functors is defined by reversing the arrows; the satellites $\{S_i F\}$ are then written as $\{S^{-i} F\}$ and called the **left satellites**.

Let \mathcal{C} be an Abelian category with enough injectives. An **injective resolution** of an object A is a right resolution $Q = \{Q^i\}$ such that all Q^i are injective. Every A admits an injective resolution, which is unique up to chain equivalence (H. Cartan, 1950). For a covariant functor $F: \mathcal{C} \rightarrow \mathcal{C}'$, the functor $A \rightarrow H^i(F(Q))$, called the i th **right derived functor** $R^i F$ of F , is independent of Q . $\{R^i F\}$ is a cohomological functor. By the universality of satellites, there exists a morphism of ∂ -functors $\{S^i F\} \rightarrow \{R^i F\}$ which is an isomorphism if and only if F is left exact. The

left derived functors $L_i F$ of a contravariant functor F are defined similarly and are isomorphic to the left satellites when F is right exact. If \mathcal{C} has enough projectives (instead of injectives), we define left (right) derived functors of covariant (contravariant) functors via **projective resolutions**. For a multifunctor, we define **partial derived functors** as well as (total) derived functors of the functor viewed as a functor defined in the \dagger product category. For instance, let $T(A, B): \mathcal{C}_1 \times \mathcal{C}_2 \rightarrow \mathcal{C}'$ be contravariant in A and covariant in B . When \mathcal{C}_2 has enough injectives, we obtain $R_2^i T(A, B) = H^i(T(A, Q))$ using an injective resolution Q of B . Suppose that T satisfies condition (i) $A \rightarrow T(A, B)$ is exact for any injective B . Then for a fixed injective B , $R_2^i T(A, B)$ is a cohomological functor in A . When A has a projective resolution P in \mathcal{C}_1 , we obtain $R_1^i T(A, B) = H^i(T(P, B))$ as well as the equation for the total derived functor $R^i T(A, B) = H^i(T(P, Q))$. We say that a functor T is **right balanced** if it satisfies (i) and also (ii) $B \rightarrow T(A, B)$ is exact for any projective A . In this case, the three derived functors are isomorphic. The **left balanced** functors are defined similarly. When the right derived functors of the functor Hom (which defines the category) exist, they are denoted by $\text{Ext}^i(A, B)$.

J. Spectral Sequences

In this section, we deal with cohomology in the category ${}_R\mathcal{M}$ of R -modules. A similar theory for homology is obtained by modifying the theory in a natural way. Similar constructions are also possible for general Abelian categories [3, 8].

A **filtration** F of a module A is a family of submodules $\{F^p(A) | p \in \mathbb{Z}\}$ such that $F^p(A) \supset F^{p+1}(A)$. We say that the filtration F is **convergent from above** (or **exhaustive**) if $\bigcup_p F^p(A) = A$, and F is **bounded from below** (or **discrete**) if $F^p(A) = 0$ for some p . The \dagger graded module $G(A) = \{G^p(A) = F^p(A)/F^{p+1}(A) | p \in \mathbb{Z}\}$ is said to be **associated** with A . A **morphism** of filtered modules $f: A \rightarrow A'$ is a module homomorphism such that $f(F^p(A)) \subset F^p(A')$. It induces a homomorphism of the graded modules $G(A) \rightarrow G(A')$. A filtration of a complex $C = \{C^n, d\}$ consists of subcomplexes $F^p(C) = \{F^p(C^n)\}$ such that $F^p(C) \supset F^{p+1}(C)$. We assume that the complex C satisfies $\bigcup_p F^p(C) = C$, and is bounded from below; i.e., for every n there exists some p such that $F^p(C^n) = 0$. In particular, if $F^0(C) = C$, $F^{p+1}(C^p) = 0$, the complex C is called **canonically bounded**. Writing $C^{p,q} = G^p(C^{p+q})$, we obtain a \dagger bigraded module $\{C^{p,q}\}$, in which p , q , and $p+q$ are called the

filtration degree, the **complementary degree**, and the **total degree**, respectively.

A **spectral sequence** $\{E_r\}$ with a graded module $D = \{D^n\}$ as its **limit** (denoted by $E_2^{p,q} \Rightarrow_p D^n$) consists of a family of doubly graded modules $E_r = \{E_r^{p,q} \mid p, q \in \mathbb{Z}\}$ ($r \geq 2$ or sometimes $r \geq 1$) and differentiations $d_r: E_r^{p,q} \rightarrow E_r^{p+r, q-r+1}$ ($p, q \in \mathbb{Z}$) of degree $(r, 1-r)$ satisfying $d_r^2 = 0$ and satisfying the following two conditions: (i) $H(E_r)$ (with respect to d_r) is isomorphic to E_{r+1} (hence there exists a sequence of graded submodules of $E_2: 0 = B_2 \subset B_3 \subset \dots \subset Z_3 \subset Z_2 = E_2$ such that $Z_r/B_r \cong E_r$); and (ii) there are submodules Z_∞ and B_∞ such that $\bigcup_k B_k \subset B_\infty \subset Z_\infty \subset \bigcap_k Z_k$, and $E_\infty = Z_\infty/B_\infty$ is isomorphic to the doubly graded module associated with a certain filtration F of D (i.e., $E_\infty^{p,q} \cong G^p(D^{p+q})$). We assume that $Z_\infty = \bigcap_k Z_k$ and $B_\infty = \bigcup_k B_k$ (**weak convergence**). Suppose that F is convergent from above and bounded from below and that $Z_k(E_2^{p,q})$ is stationary for every p, q . Then $\{E_r\}$ is called **regular**. $\{E_r\}$ is **bounded from below** if for every n there exists a p_0 such that $E_2^{p,q} = 0$ for $p < p_0$. In particular, if $E_2^{p,q} = 0$ ($p < 0, q < 0$), then $\{E_r\}$ is called the **first quadrant** (or **cohomology spectral sequence**). In the latter case, the **edge homomorphisms** $E_2^{p,0} \rightarrow E_\infty^{p,0}$, $E_\infty^{0,q} \rightarrow E_2^{0,q}$ are defined through **base terms** $E_k^{p,0}$ and **fiber terms** $E_k^{0,q}$, respectively. A morphism of spectral sequences $f: \{E_r, D\} \rightarrow \{E'_r, D'\}$ consists of $f_r: E_r \rightarrow E'_r$ of degree $(0, 0)$ and $f: D \rightarrow D'$ of degree 0 which preserve the mechanism of spectral sequences. When the spectral sequences are regular, a morphism f is an isomorphism if one of the f_r is an isomorphism. Addition is naturally introduced in the set of morphisms so that spectral sequences form an additive category. An additive functor from an Abelian category \mathcal{C} to this category is called a **spectral functor**. A filtered complex $\{C, F\}$ gives rise to a spectral sequence $E_2^{p,q} \Rightarrow G(H(C))$ if we put $Z_r^p = \{a \in F^p(C) \mid da \in F^{p+r}(C)\}$, $B_r^p = dZ_r^{p-r}$, $E_r^p = Z_r^p/(Z_{r-1}^{p+1} + B_{r-1}^p)$, $E_r = \sum_p E_r^p$. A double complex $C = \{C^{p,q}, d_I, d_{II}\}$ admits two natural filtrations $F_I: F_I^p(C) = \sum_{s \geq p} \sum_q C^{s,q}$ and $F_{II}: F_{II}^q(C) = \sum_{t \geq q} \sum_p C^{p,t}$. By the procedure above, these filtrations give rise to spectral sequences $H_I^p(H_{II}^q(C)) \Rightarrow_p H^n(C)$ and $H_{II}^q(H_I^p(C)) \Rightarrow_q H^n(C)$, respectively. Comparison of these sequences yields many useful results. Let T be an additive covariant functor from an Abelian category \mathcal{C} to ${}_R\mathcal{M}$, C be a complex in \mathcal{C} , and $Q = \{Q^{p,q}\}$ be an injective resolution of C . The double complex Q gives rise to spectral sequences $H^p(R^q T(C)) \Rightarrow H(T(Q))$ and $R^p T(H^q(C)) \Rightarrow H(T(Q))$. The limit $H(T(Q))$ is independent of Q and is called the **hypercohomology** of T with respect to C [2, 8]. We can similarly define hypercohomology of multifunctors. The theory of

spectral sequences was initiated by J. Leray (1946), and suitable algebraic formulations were given by J. L. Koszul (1950).

K. Categories of Modules

The category ${}_R\mathcal{M}$ (resp. \mathcal{M}_R) of left (right) R -modules over a \dagger unitary ring R is not only an Abelian category but also a Grothendieck category (\rightarrow 277 Modules). The \dagger full embedding theorem permits us to deduce many propositions about general Abelian categories from the consideration of ${}_R\mathcal{M}$. An object P of ${}_R\mathcal{M}$ is projective if and only if it is isomorphic to a direct summand of a \dagger free module. Any projective module is the direct sum of countably generated projective modules (I. Kaplansky, 1958). It follows that any projective module over a \dagger local ring is free. Finitely generated projective modules P_1 and P_2 are said to be equivalent if there exist finitely generated free modules F_1, F_2 such that $P_1 \oplus F_1 \cong P_2 \oplus F_2$. The equivalence classes then form an Abelian group with respect to the direct sum construction called the **projective class group** of a ring R . The category of complex \dagger vector bundles over a compact space X is equivalent to the category of projective modules over $C(X)$, the ring of complex-valued continuous functions on X , and similarly for other types of spaces and bundles. Many investigations have been made involving the problem of whether every projective module over a polynomial ring is free (J.-P. Serre, 1955). This problem was settled affirmatively by D. Quillen [13] and independently by A. Suslin. It has been observed that "big" projective modules are often free: for example, nonfinitely generated projective modules over an \dagger indecomposable weakly Noetherian ring are free (Y. Hinohara, 1963).

The n th right derived functor of $\text{Hom}_R(A, B)$ is denoted by $\text{Ext}_R^n(A, B)$ (\rightarrow Section G). This is a bifunctor ${}_R\mathcal{M} \times {}_R\mathcal{M} \rightarrow (\text{Ab})$, contravariant in A and covariant in B . Ext_R^0 is isomorphic to and identified with Hom_R . An exact sequence $0 \rightarrow A' \rightarrow A \rightarrow A'' \rightarrow 0$ gives rise to the connecting homomorphisms $\Delta^n: \text{Ext}_R^n(A', B) \rightarrow \text{Ext}_R^{n+1}(A'', B)$, and the following sequence is exact: $\dots \rightarrow \text{Ext}_R^{n-1}(A', B) \xrightarrow{\Delta} \text{Ext}_R^n(A', B) \rightarrow \text{Ext}_R^n(A, B) \rightarrow \text{Ext}_R^n(A'', B) \xrightarrow{\Delta} \text{Ext}_R^{n+1}(A'', B) \rightarrow \dots$ (the exact sequence of Ext). Similarly, an exact sequence $0 \rightarrow B' \rightarrow B \rightarrow B'' \rightarrow 0$ gives rise to $\Delta^n: \text{Ext}_R^n(A, B') \rightarrow \text{Ext}_R^{n+1}(A, B'')$ and to an exact sequence of Ext . An **extension** of A by B (or of B by A) is an exact sequence $(E): 0 \rightarrow B \rightarrow X \rightarrow A \rightarrow 0$. The set of equivalence classes of extensions of A by B is in one-to-one correspondence with $\text{Ext}_R^1(A, B)$ by assigning to (E) its **characteristic class** $\chi_E = \Delta^0(1) \in \text{Ext}_R^1(A, B)$,

where 1 denotes the identity of $\text{Hom}_R(B, B)$. In this correspondence, the sum of two extensions is obtained by a construction called **Baer's sum** of extensions. Similarly, $\text{Ext}_R^n(A, B)$ is interpreted as the set of the equivalence classes of n -fold extensions $0 \rightarrow B \rightarrow X_{n-1} \rightarrow \dots \rightarrow X_0 \rightarrow A \rightarrow 0$ (exact). This point of view permits us to establish a theory of Ext , etc., in more general (additive) categories lacking enough projectives or injectives (N. Yoneda, 1954, 1960).

The tensor product $A \otimes_R B$ is a right exact covariant bifunctor $\mathcal{M}_R \otimes_R \mathcal{M} \rightarrow (\text{Ab})$. If the functor $t_P(\cdot) = \cdot \otimes P$ is exact, P is called a \dagger flat module. A projective module is flat. In general, a flat module is the \dagger inductive limit of finitely generated free modules (M. Lazard, 1964). A flat module P is called \dagger faithfully flat if $P \neq \mathfrak{m}P$ for every maximal ideal \mathfrak{m} of R . The functors \otimes and Hom are related by \dagger adjointness (\rightarrow 52 Categories and Functors). From this viewpoint \otimes can be introduced in more general categories. Left-derived functors of $A \otimes_R B$ are denoted by $\text{Tor}_n^R(A, B)$ and are called n th **torsion products** of A and B . $\text{Tor}_1^R(A, B)$ is often denoted by $A * _R B$. The functor \otimes_R is left balanced, hence Tor is calculated by using projective resolutions of A , B , or both A and B . We have $\text{Tor}_0^R = \otimes_R$. An exact sequence $0 \rightarrow A' \rightarrow A \rightarrow A'' \rightarrow 0$ gives rise to $\Delta_n: \text{Tor}_{n+1}^R(A'', B) \rightarrow \text{Tor}_n^R(A', B)$ and the infinite exact sequence of Tor , and similarly for the second variables.

From the various relations between Hom and \otimes follow the corresponding relations between their derived functors. When Λ and Γ are algebras over K and $\Omega = \Lambda \otimes \Gamma$, we can define the **external product** (\top -**product**), which is a mapping

$$\begin{aligned} \top: \text{Tor}_p^\Lambda(A, B) \otimes \text{Tor}_q^\Gamma(A', B') \\ \rightarrow \text{Tor}_{p+q}^\Omega(A \otimes A', B \otimes B'). \end{aligned}$$

In particular, if Λ and Γ are K -projective and $\text{Tor}_n^K(A, A') = 0$ ($n > 0$), then we can define the **wedge product** (\vee -**product**) $\vee: \text{Ext}_\Lambda^n(A, B) \otimes \text{Ext}_\Gamma^q(A', B') \rightarrow \text{Ext}_\Omega^{n+q}(A \otimes A', B \otimes B')$. The latter is described in terms of the composition of module extensions. When $K = \Lambda = \Gamma = \Omega$, the \top -product reduces to the **internal product**, called the \mathfrak{m} -**product**. If Λ is a \dagger Hopf algebra over K , the \dagger comultiplication $\Lambda \rightarrow \Lambda \otimes \Lambda$ induces $\text{Ext}_{\Lambda \otimes \Lambda} \rightarrow \text{Ext}_\Lambda$. This, combined with the \vee -product, yields the **cup product** (\smile -**product**) $\smile: \text{Ext}_\Lambda^n(A, B) \otimes \text{Ext}_\Lambda^q(A', B') \rightarrow \text{Ext}_\Lambda^{n+q}(A \otimes A', B \otimes B')$. We define similarly \perp -**product**, \wedge -**product**, ω -**product**, and \frown -**product** (**cap product**) [2]. Let Λ , Γ , and Σ be algebras over K , with Λ K -projective; let $A \in \mathcal{M}_{\Lambda \otimes \Gamma}$, $B \in \mathcal{M}_{\Lambda \otimes \Sigma}$, $C \in \mathcal{M}_{\Gamma \otimes \Sigma}$, and assume $\text{Tor}_n^\Lambda(A, B) = 0$ ($n > 0$). The natural isomorphism $\text{Hom}_{\Lambda \otimes \Gamma}(A, \text{Hom}_\Sigma(B, C)) \cong \text{Hom}_{\Gamma \otimes \Sigma}(A \otimes_\Lambda B, C)$ then yields a spectral sequence

$\text{Ext}_{\Lambda \otimes \Gamma}^p(A, \text{Ext}_\Sigma^q(B, C)) \Rightarrow_p \text{Ext}_{\Gamma \otimes \Sigma}^n(A \otimes_\Lambda B, C)$ by the double complex argument in Section D.

The **homological dimension** $\text{hdim}_R A$, $\text{dh}_R A$, or **projective dimension** $\text{proj dim}_R A$ of $A \in {}_R \mathcal{M}$ is the supremum ($\leq \infty$) of n such that $\text{Ext}_R^n(A, B) \neq 0$ for some B . The relation $\text{hdim}_R A \leq 0$ means that A is projective. The **injective dimension** $\text{inj dim}_R B$ of $B \in {}_R \mathcal{M}$ is defined similarly by means of the functor $\text{Ext}_R^n(\cdot, B)$, and the **weak dimension** $\text{wdim}_R C$ of $C \in {}_R \mathcal{M}$ by the functor $\text{Tor}_n^R(\cdot, C)$. The common value $\sup\{\text{proj dim}_R A \mid A \in {}_R \mathcal{M}\} = \sup\{\text{inj dim}_R B \mid B \in {}_R \mathcal{M}\}$ is called the **left global dimension** $1 \text{ gl dim } R$ of R . It is identical with the supremum of homological dimensions of \dagger cyclic modules (M. Auslander, 1955). The **right global dimension** $\text{r gl dim } R$ is defined similarly. The common value $\sup\{\text{wdim}_R A \mid A \in \mathcal{M}_R\} = \sup\{\text{wdim}_R C \mid C \in {}_R \mathcal{M}\}$ is called the **weak global dimension** $\text{w gl dim } R$ of R . We have $\text{w gl dim } R \leq 1 \text{ gl dim } R$, $\text{r gl dim } R$. The equality may fail to hold (Kaplansky, 1958). If R is \dagger Noetherian, the three global dimensions coincide (Auslander, 1955) and are called simply the **global dimension** of R : $\text{gl dim } R$. The condition $1 \text{ gl dim } R = 0$ (or $\text{r gl dim } R = 0$) holds if and only if R is an \dagger Artinian semi-simple ring, while $\text{w gl dim } R = 0$ if and only if R is a \dagger regular ring in the sense of J. von Neumann (M. Harada, 1956). A ring R is called **left (right) hereditary** if $1 \text{ gl dim } R \leq 1$ ($\text{r gl dim } R \leq 1$), and **left (right) semihereditary** if every finitely generated left (right) ideal is projective. A left and right (semi)hereditary ring is called a **(semi)hereditary ring**. Since projectivity and \dagger invertibility of an ideal of a (commutative) integral domain R are equivalent, R is hereditary if and only if R is a \dagger Dedekind ring. In this case, the projective class group of R reduces to the \dagger ideal class group. An integral domain R is semihereditary if and only if $\text{w gl dim } R \leq 1$ (A. Hattori, 1957), and in that case R is called a **Prüfer ring**. A \dagger maximal order over a Dedekind ring is hereditary. A commutative semihereditary ring R is characterized by the property that flatness of R -modules is equivalent to torsion-freeness (S. Endo, 1961). A Noetherian ring R is left self-injective if and only if R is \dagger quasi-Frobenius (M. Ikeda, 1952), and the global dimension of a quasi-Frobenius ring is 0 or ∞ (S. Eilenberg and T. Nakayama, 1955). A polynomial ring $R = K[X_1, \dots, X_n]$ over a commutative ring K satisfies $\text{gl dim } R = \text{gl dim } K + n$. When K is a field, this is a reformulation of Hilbert's theory of syzygy sequences (\rightarrow 369 Rings of Polynomials). In this sense, the study of the global dimension of rings and categories is sometimes called **syzygy theory** (Eilenberg, 1956). The homological algebra of commutative Noetherian rings has been studied extensively and

is useful in algebraic geometry. Since $\text{gl dim } R = \sup_m \text{gl dim } R_m$ (R_m is the \dagger ring of quotients relative to m), with m running over the maximal ideals of R , the problem of determining $\text{gl dim } R$ reduces to the case of \dagger local rings. A finitely generated flat module over a local ring R is free. If K denotes the residue field R/m , where m is the maximal ideal of the local ring R , $\text{Tor}^R(K, K)$ has the structure of a Hopf algebra (E. F. Assmus, Jr., 1959). Detailed results concerning the **Betti numbers** $\dim \text{Tor}_i^R(K, K)$ of R have been obtained (J. Tate, 1957, and others). In particular, R is \dagger regular if and only if $\text{gl dim } R < \infty$ (Serre, 1955). A local ring R is called a **Gorenstein ring** if the injective dimension of R -module R is finite. This is a notion intermediate between regular rings and \dagger Macaulay rings (\rightarrow 284 Noetherian Rings).

Consideration of a ring R in relation to a subring S leads to **relative homological algebra**. Foundations for this theory were established by Hochschild (1956). An exact sequence of R -modules that \dagger splits as a sequence of S -modules is called an **(R, S) -exact sequence**. An R -module P is called an **(R, S) -projective module** if $\text{Hom}_R(P, \cdot)$ maps any (R, S) -exact sequence to an exact sequence. **(R, S) -injective modules** are defined similarly. Based on these notions, $\text{Ext}_{(R, S)}^i$ and $\text{Tor}^{(R, S)}$ are defined as the **relative derived functors** of Hom_R and \otimes_R , respectively. We also have a relative theory from a different viewpoint (S. Takasu, 1957). Relative theory is extended to general categories from various viewpoints [6, 14] (\rightarrow Section Q).

L. Cohomology Theory for Associative Algebras

Let Λ be an \dagger algebra over a commutative ring K and A a \dagger two-sided Λ -module. Let C^n be the module of all n -linear mappings of Λ into A called **n -cochains** ($C^0 = A$). Define the coboundary operator $\delta^n: C^n \rightarrow C^{n+1}$ by $(\delta^n f)(\lambda_1, \dots, \lambda_{n+1}) = \lambda_1 f(\lambda_2, \dots, \lambda_{n+1}) + \sum_{i=1}^n (-1)^i f(\lambda_1, \dots, \lambda_i \lambda_{i+1}, \dots, \lambda_{n+1}) + (-1)^{n+1} f(\lambda_1, \dots, \lambda_n) \lambda_{n+1}$. We thus obtain a complex whose cohomology is denoted by $H^n(\Lambda, A)$ and is called the **n th Hochschild's cohomology group** of Λ relative to the **coefficient module** A (Hochschild, 1945). A cochain f is called **normalized** if $f(\lambda_1, \dots, \lambda_n) = 0$ whenever one of the λ_i is 1. We obtain the same cohomology group $H^n(\Lambda, A)$ from the subcomplex of normalized cochains. $\{H^n(\Lambda, \cdot)\}$ is a cohomological functor from the category ${}_{\Lambda}\mathcal{M}_{\Lambda}$ of two-sided Λ -modules to the category ${}_K\mathcal{M}$. Using the **enveloping algebra** $\Lambda^e = \Lambda \otimes_K \Lambda^0$, where Λ^0 is an anti-isomorphic

copy of Λ , ${}_{\Lambda}\mathcal{M}_{\Lambda}$ can be identified with ${}_{\Lambda^e}\mathcal{M}$ (and \mathcal{M}_{Λ^e}). If Λ is K -projective, $\{H^n(\Lambda, \cdot)\}$ is isomorphic to $\{\text{Ext}_{\Lambda^e}^n(\Lambda, \cdot)\}$. (In [2], $H^n(\Lambda, A)$ is defined as $\text{Ext}_{\Lambda^e}^n(\Lambda, A)$ in general.) We have $H^0(\Lambda, A) = \{a \in A \mid \lambda a = a\lambda, \forall \lambda \in \Lambda\}$. We call 1-cocycles **derivations** (or **crossed homomorphisms**) of Λ in A and 1-coboundaries **inner derivations**. Thus $H^1(\Lambda, A)$ is the derivation class group and is related to the \dagger ramification. When K is a field, $H^1(\Lambda, \cdot) = 0$ if and only if Λ is a \dagger separable algebra. In general, an algebra Λ over a commutative ring K is called a **separable algebra** if Λ is Λ^e -projective, i.e., if $\text{Ext}_{\Lambda^e}^1(\Lambda, \cdot) = 0$ (Auslander and O. Goldman, 1960). This is a generalization of the notion of \dagger maximally central algebras (Nakayama and G. Azumaya, 1948). We have a one-to-one correspondence of $H^2(\Lambda, A)$ to the family of **algebra extensions** of Λ with kernel A (i.e., K -algebras Γ containing A as a two-sided ideal such that $\Gamma/A = \Lambda$) satisfying $A^2 = 0$. Any extension of an algebra Λ over a field K such that $H^2(\Lambda, \cdot) = 0$ splits over a nilpotent kernel (J. H. C. Whitehead and G. Hochschild). This holds in particular for a separable algebra, and we obtain the \dagger Wedderburn-Mal'tsev theorem. There are some interpretations of $H^3(\Lambda, A)$ in terms of extensions.

The supremum (\leq) of n such that $H^n(\Lambda, A) \neq 0$ for some A is called the **cohomological dimension** of Λ and written $\dim \Lambda$. For a finite-dimensional algebra Λ over a field K , $\dim \Lambda < \infty$ if and only if Λ/N is separable and $\text{gl dim } \Lambda < \infty$, where N is the \dagger radical of Λ (N. Ikeda, H. Nagao, and Nakayama, 1954).

The homology groups $H_n(\Lambda, A)$ of Λ relative to a coefficient module A are defined similarly. If Λ is K -projective, $\{H_n(\Lambda, \cdot)\}$ is isomorphic to $\{\text{Tor}_n^{\Lambda^e}(\cdot, \Lambda)\}$.

M. Cohomology of Groups

The pair consisting of an algebra Λ over K and an algebra homomorphism $\varepsilon: \Lambda \rightarrow K$ is called a **supplemented algebra** [2] (or **augmented algebra** [6]), of which ε is the **augmentation**. The \dagger group algebra $\mathbf{Z}[G]$ of a group G over the ring of rational integers is a supplemented algebra, in which the augmentation is defined by $\varepsilon(x) = 1$ ($x \in G$). The category of left G -modules is identified with the category of left $\mathbf{Z}[G]$ -modules. For a finite group G , a finitely generated projective G -module is not necessarily free (D. S. Rim, 1959) and is isomorphic to the direct sum of a free module and a left ideal of $\mathbf{Z}[G]$. It follows that the projective class group of $\mathbf{Z}[G]$ is a finite group (R. G. Swan, 1960). The **cohomology groups** and **homology groups** of G relative to $A \in {}_G\mathcal{M}$ (Eilenberg and S. MacLane, 1943) are defined

by $H^n(G, A) = \text{Ext}_{\mathbb{Z}[G]}^n(\mathbb{Z}, A)$ and $H_n(G, A) = \text{Tor}_{\mathbb{Z}[G]}^n(\mathbb{Z}, A)$, respectively. Their concrete description is given usually via the $\mathbb{Z}[G]$ -standard resolution of \mathbb{Z} .

(1) **Homogeneous formulation.** The group of **homogeneous n -chains** is the free Abelian group with basis $G \times \dots \times G$ ($n+1$ times), on which G operates by $x(x_0, \dots, x_n) = (xx_0, \dots, xx_n)$, and the boundary operator is defined by $d(x_0, \dots, x_n) = \sum_{i=0}^n (-1)^i (x_0, \dots, \hat{x}_i, \dots, x_n)$.

(2) **Nonhomogeneous formulation.** The group of **nonhomogeneous n -chains** is the $\mathbb{Z}[G]$ -free module with basis $G \times \dots \times G$ (n times), and the boundary operator is defined by $d(x_1, \dots, x_n) = x_1(x_2, \dots, x_n) + \sum_{i=1}^n (-1)^i (x_1, \dots, x_i x_{i+1}, \dots, x_n) + (-1)^n (x_1, \dots, x_{n-1})$. A nonhomogeneous 2-cocycle is sometimes called a **factor set**. $H^0(G, A)$ is the submodule A^G of A consisting of the G -invariant elements, while $H_0(G, A)$ is the largest residue class module A_G of A on which G acts trivially. Given two groups G and K , an exact sequence of group homomorphisms $1 \rightarrow K \rightarrow E \rightarrow G \rightarrow 1$ is called a **group extension** of G over the kernel K . When K is Abelian, the extension canonically induces a G -module structure on K , and the deviation of K from being a semidirect factor of E is measured by a factor set. The group $H^2(G, A)$ is thus in one-to-one correspondence with the set of equivalence classes of the \dagger group extensions of G over A which induce the originally given G -module structure on A (\rightarrow 190 Groups N). This point of view is essential in the proof of the \dagger Schur-Zassenhaus theorem (\rightarrow 151 Finite Groups). $H^3(G, A)$ is interpreted as the set of obstructions for extensions (Eilenberg and MacLane, 1947). For a \dagger free group F , $H^n(F, A) = 0$ ($n > 1$). If a group G is represented as a factor group F/R of a free group F , we have a group extension $1 \rightarrow K \rightarrow E \rightarrow G \rightarrow 1$, where $K = R/[R, R]$ and $E = F/[R, R]$. Let $\xi \in H^2(G, K)$ correspond to this extension. Then for any G -module A , the cup product $\chi \rightarrow \chi \smile \xi$ followed by the pairing $\text{Hom}(K, A) \otimes K \rightarrow A$ provides isomorphisms $H^n(G, \text{Hom}(K, A)) \cong H^{n+2}(G, A)$ ($n > 0$) (the **cup product reduction theorem** of Eilenberg and MacLane, 1947); similarly, we have the reduction theorem for the homology. The \mathbb{Z} -algebra $H(G, \mathbb{Z}) = \sum_{n=0}^{\infty} H^n(G, \mathbb{Z})$ under the multiplication defined by the cup product is finitely generated if G is a finite group (B. B. Venkov, 1959; L. Evens, 1961).

The following are mappings relative to a subgroup H .

(1) The inner automorphism by $x \in G$ induces an isomorphism of $H^n(H, A)$ and $H^n(xHx^{-1}, A)$ which reduces to the identity of $H^n(G, A)$ if $H = G$. Hence if H is a normal subgroup, $H^n(H, A)$ has the structure of a G/H -module, and similarly for $H_n(H, A)$.

(2) When H is a normal subgroup of G , the mapping $(x_1, \dots, x_n) \rightarrow (x_1 H, \dots, x_n H)$ of nonhomogeneous chains induces the **inflation** (or **lift**) $\text{Inf}: H^n(G/H, A^H) \rightarrow H^n(G, A)$ and the **deflation** $\text{Def}: H_n(G, A) \rightarrow H_n(G/H, A_H)$.

(3) The embedding of nonhomogeneous chains induces the **restriction** $\text{Res}: H^n(G, A) \rightarrow H^n(H, A)$ and the **injection** Inj (or **corestriction** Cor): $H_n(H, A) \rightarrow H_n(G, A)$. The theory of \dagger induced representation gives another construction of these mappings; that is, if we put $\iota^G(A) = \text{Hom}_{\mathbb{Z}[H]}(\mathbb{Z}[G], A)$, Res is obtained by the isomorphism $H^n(G, \iota^G(A)) \cong H^n(H, A)$ combined with the homomorphism induced by $A \rightarrow \iota^G(A)$; while if we put $\iota_G(A) = \mathbb{Z}[G] \otimes_{\mathbb{Z}[H]} A$, Inj is obtained by the isomorphism $H_n(H, A) \cong H_n(G, \iota_G(A))$ followed by the homomorphism induced by $\iota_G(A) \rightarrow A$.

(4) If $(G: H) < \infty$, we have $\iota^G(A) \cong \iota_G(A)$. The composition of $H^n(H, A) \rightarrow H^n(G, \iota_G(A)) \rightarrow H^n(G, A)$ defines on the cohomology groups the $\text{Inj}: H^n(H, A) \rightarrow H^n(G, A)$, while the composition $H_n(G, A) \rightarrow H_n(G, \iota^G(A)) \rightarrow H_n(H, A)$ gives the $\text{Res}: H_n(G, A) \rightarrow H_n(H, A)$. In particular, $\text{Res}: H_1(G, \mathbb{Z}) \rightarrow H_1(H, \mathbb{Z})$ coincides with the \dagger transfer $G/[G, G] \rightarrow H/[H, H]$.

(5) Let H be a normal subgroup of G . Consider the additive relation ρ (the correspondence) between $h \in Z^n(H, A)^G$ and $f \in Z^{n+1}(G/H, A^H)$ determined by $\rho(h, f)$ if and only if there exists a $g \in C^n(G, A)$ such that $h = \text{Res } g$ and $\text{Inf } f = \delta g$. If the relation induces a homomorphism $H^n(H, A)^G \rightarrow H^{n+1}(G/H, A^H)$, then it is called the **transgression**. If $H^i(H, A) = 0$ ($0 < i < n$), the sequence $0 \rightarrow H^n(G/H, A^H) \rightarrow H^n(G, A) \rightarrow H^n(H, A)^G \rightarrow H^{n+1}(G/H, A^H) \rightarrow H^{n+1}(G, A)$ composed of inflation, restriction, and transgression mappings is exact (Hochschild and Serre, 1953) and is called the **fundamental exact sequence**. This exact sequence can be derived from a certain spectral sequence $H^n(G/H, H^q(H, A)) \Rightarrow_p H^n(G, A)$ (R. C. Lyndon, 1948; Hochschild and Serre, 1953).

The relative (co)homology theory relative to a subgroup (I. T. Adamson, 1954) can be dealt with in terms of the relative Ext and the relative Tor (Hochschild, 1956). Many results in the absolute case are generalized to the relative case: for example, the fundamental exact sequence (Nakayama and Hattori, 1958). The relative theory is further generalized to the cohomology theory of \dagger permutation representations of G (E. Snapper, 1964).

Non-Abelian Cohomology. For a non-Abelian G -group A , the cohomology "set" $H^1(G, A)$ (and $H^0(G, A)$) is defined as in the Abelian case by means of the nonhomogeneous cochains (\rightarrow e.g., [9]). Some efforts are being made toward the construction of a more general non-

Abelian theory (J. Giraud, *Cohomologie non abélienne*, Springer, 1971).

N. Finite Groups

Let G be a finite group and A a G -module. Define the norm $N: A \rightarrow A$ by $N(a) = \sum_{x \in G} xa$, and denote $\text{Ker } N$ by ${}_N A$. The kernel of the augmentation $\varepsilon: \mathbb{Z}[G] \rightarrow \mathbb{Z}$ is denoted by I . Put $\hat{H}^n(G, A) = H^n(G, A)$ ($n > 0$), $\hat{H}^0(G, A) = A^G / NA$, $\hat{H}^{-1}(G, A) = {}_N A / IA$, and $\hat{H}^{-n}(G, A) = H_{n-1}(G, A)$ ($n > 1$). Then $\{\hat{H}^n(G, \cdot)\}$ forms a cohomological functor, called the **Tate cohomology** (E. Artin and J. T. Tate), that can be described as the set of cohomology groups concerning a certain complex called a **complete free resolution** of \mathbb{Z} . (Similar arguments are valid more generally for quasi-Frobenius rings (Nakayama, 1957), and a theory of this kind is called **complete cohomology theory**.) We have $\hat{H}^n(G, A) = 0$ ($n \in \mathbb{Z}$) if and only if $\text{h dim}_{\mathbb{Z}[G]} A \leq 1$ (Nakayama, 1957). If A satisfies the conditions (i) $\hat{H}^1(G_p, A) = 0$ for any Sylow p -subgroup G_p of G , and (ii) there exists a $\xi \in \hat{H}^2(G, A)$ such that $\text{Res } \xi \in \hat{H}^2(G_p, A)$ has the same order as G_p and generates all of $\hat{H}^2(G_p, A)$, then the homomorphisms $\hat{H}^n(H, B) \rightarrow \hat{H}^{n+2}(H, A \otimes B)$ ($n \in \mathbb{Z}$) defined by the cup product with $\text{Res } \xi$ are isomorphisms for every subgroup H and every G -module B such that $\text{Tor}(A, B) = 0$ (Nakayama, 1957; for $B = \mathbb{Z}$, Tate, 1952). If G is cyclic, the mappings $\hat{H}^n(A) \rightarrow \hat{H}^{n+2}(A)$ ($n \in \mathbb{Z}$) defined by the cup product with a generator of $\hat{H}^2(\mathbb{Z})$ are isomorphisms. (The notation is abbreviated by omitting G .) If the orders of $\hat{H}^0(A)$ and $\hat{H}^1(A)$ are finite, their ratio is called the **Herbrand quotient** $h(A)$ of A . If $0 \rightarrow A' \rightarrow A \rightarrow A'' \rightarrow 0$ is exact, then $h(A) = h(A')h(A'')$. If A is finite, then $h(A) = 1$. By combining these two facts we obtain **Herbrand's lemma**: If A' is a G -submodule of A of finite index and $h(A')$ exists, then $h(A)$ also exists and $h(A) = h(A')$. The periodicity $\hat{H}^n(A) = \hat{H}^{n+p}(A)$ ($n \in \mathbb{Z}$, $A \in {}_G \mathcal{M}$) holds if and only if every Sylow subgroup is cyclic or a \dagger generalized quaternion group (Artin and Tate; [2]).

Let L/K be a finite \dagger Galois extension with the \dagger Galois group G . The cohomology groups of various types of G -modules related to L/K are called the **Galois cohomology groups** (\rightarrow 172 Galois Theory). Using **continuous cocycles**, a cohomology theory (**Tate cohomology**) is developed for infinite Galois extensions as well [9, 10]. By means of Galois cohomology (\rightarrow 59 Class Field Theory), the cohomology theory of finite groups and of \dagger totally disconnected compact groups (which are \dagger profinite groups) plays an important role in class field theory and its related areas.

O. Cohomology Theory of Lie Algebras

Let \mathfrak{g} be a \dagger Lie algebra over a commutative ring K , and assume that \mathfrak{g} is K -free. The \dagger enveloping algebra $U = U(\mathfrak{g})$ is a \dagger supplemented algebra over K . For a \mathfrak{g} -module ($= U$ -module) A , $\text{Ext}_U^n(K, A)$ and $\text{Tor}_n^U(K, A)$ are called the **cohomology groups** $H^n(\mathfrak{g}, A)$ and **homology groups** $H_n(\mathfrak{g}, A)$, respectively, of \mathfrak{g} relative to the coefficient module A . They are usually described by means of the U -free resolution $U \otimes \wedge_K(\mathfrak{g})$ of K (called the **standard complex** of \mathfrak{g}) constructed by C. Chevalley and Eilenberg (1948), where $\wedge_K(\mathfrak{g})$ is the exterior algebra of the K -module \mathfrak{g} and (denoting $1 \otimes (x_1 \wedge \dots \wedge x_n)$ by (x_1, \dots, x_n)) the differentiation is given by

$$d(x_1, \dots, x_n) = \sum_{i=1}^n (-1)^{i+1} x_i(x_1, \dots, \hat{x}_i, \dots, x_n) + \sum_{1 \leq i < j \leq n} (-1)^{i+j} ([x_i, x_j], x_1, \dots, \hat{x}_i, \dots, \hat{x}_j, \dots, x_n).$$

For $n > [\mathfrak{g}: K]$, $H^n(\mathfrak{g}, A) = H_n(\mathfrak{g}, A) = 0$. If \mathfrak{g} is a \dagger semisimple Lie algebra over a field K of characteristic 0, we have $H^1(\mathfrak{g}, A) = 0$, $H^2(\mathfrak{g}, A) = 0$, while $H^3(\mathfrak{g}, K) \neq 0$. $H^1(\mathfrak{g}, A) = 0$ is equivalent to Weyl's theorem, which asserts the complete reducibility of finite-dimensional representations (\rightarrow 248 Lie Algebras E). $H^2(\mathfrak{g}, A)$ and $H^3(\mathfrak{g}, A)$ are interpreted by means of Lie algebra extensions as in the cohomology of groups. The theorem on \dagger Levi decomposition is derived from $H^2(\mathfrak{g}, A) = 0$. Chevalley and Eilenberg constructed this cohomology by algebraization of the cohomology of compact \dagger Lie groups. They also introduced the notion of relative cohomology groups $H^n(\mathfrak{g}, \mathfrak{h}, A)$ relative to a Lie subalgebra \mathfrak{h} of \mathfrak{g} , which correspond to the cohomology of homogeneous spaces. $H^n(\mathfrak{g}, \mathfrak{h}, A)$ does not always coincide with $\text{Ext}_{U(\mathfrak{g}), U(\mathfrak{h})}^n(K, A)$ (Hochschild, 1956), but does so in an important case where K is a field of characteristic 0 and \mathfrak{h} is \dagger reductive in \mathfrak{g} (\rightarrow 248 Lie Algebras).

For \dagger transformation spaces of \dagger linear algebraic groups G over a field K , the **rational cohomology groups** are introduced using the notion of **rational injectivity** (Hochschild, 1961). In particular, if G is a \dagger unipotent algebraic group over a field K of characteristic 0, then $H(G, A)$ is isomorphic to $H(\mathfrak{g}, A)$, where \mathfrak{g} is the Lie algebra of G . There is also a relative theory.

P. Amitsur Cohomology

Let R be a commutative ring, and F a covariant functor from the category \mathcal{C}_R of com-

mutative R -algebras to the category of Abelian groups. For $S \in \mathcal{C}_R$ and $n = 0, 1, 2, \dots$, we write $S^{(n)} = S \otimes \dots \otimes S$ (n -fold tensor product over R). Let $\varepsilon_i: S^{(n+1)} \rightarrow S^{(n+2)}$ ($i = 0, 1, \dots, n+1$) be \mathcal{C}_R -morphisms defined by $\varepsilon_i(x_0 \otimes \dots \otimes x_n) = x_0 \otimes \dots \otimes x_{i-1} \otimes 1 \otimes x_i \otimes \dots \otimes x_n$. Defining $d^n: F(S^{(n+1)}) \rightarrow F(S^{(n+2)})$ by $d^n = \sum_{i=0}^{n+1} (-1)^i F(\varepsilon_i)$, we obtain a cochain complex $\{F(S^{(n+1)}), d^n\}$. This complex and its cohomology groups are called the **Amitsur complex** and the **Amitsur cohomology groups**, and are usually denoted by $C(S/R, F)$ and $H^n(S/R, F)$, respectively.

If S/R is a finite Galois extension with Galois group G , then the group $H^n(S/R, U)$ of the unit group functor U is naturally isomorphic to $H^n(G, U(S))$. If S/R is a finite purely inseparable extension, then $H^n(S/R, U) = 0$ for $n \geq 3$. The group $H^2(S/R, U)$ is related to the Brauer group $B(S/R)$ (\rightarrow 29 Associative Algebras K).

Q. Relative Theory

In the course of the development of homological algebra, it has been recognized that the notion of projective (resp. injective) resolutions should be generalized ([14, 4]; Hochschild, 1956). In the meantime, a method has been introduced that utilizes simplicial objects in order to define the derived functor of an arbitrary functor with Abelian category as its range ([15, 17]; J. Beck, 1967). As a consequence of these developments, there emerged a viewpoint, which we describe below, making it possible to unify various known definitions of (co-)homology theories that has been designed for particular cases.

Let \mathcal{A} be a category and \mathcal{P} a class of objects in \mathcal{A} . In this section, we denote the set $\text{Hom}_{\mathcal{A}}(A, B)$ of morphisms by $\mathcal{A}(A, B)$. A morphism $f: A \rightarrow B$ in \mathcal{A} is called a **\mathcal{P} -epimorphism** if the induced mapping $\mathcal{A}(P, f)$ ($= \text{Hom}(P, f): \mathcal{A}(P, A) \rightarrow \mathcal{A}(P, B)$) is surjective for any $P \in \mathcal{P}$. The class \mathcal{P} is called a **projective class** in \mathcal{A} if there exist an object $P \in \mathcal{P}$ and a \mathcal{P} -epimorphism $f: P \rightarrow A$ for each object $A \in \mathcal{A}$. To any category \mathcal{A} , we associate a preadditive category $\mathbf{Z}\mathcal{A}$, adding a zero object to \mathcal{A} if necessary: Put $\text{Ob } \mathbf{Z}\mathcal{A} = \text{Ob } \mathcal{A}$, $\mathbf{Z}\mathcal{A}(A, B) = \text{free Abelian group generated by the set } \mathcal{A}(A, B)$. \mathcal{A} is regarded as a subcategory of $\mathbf{Z}\mathcal{A}$ by the natural inclusion $J: \mathcal{A} \rightarrow \mathbf{Z}\mathcal{A}$. Any functor T from \mathcal{A} into an Abelian category \mathcal{B} has a unique additive extension $\bar{T}: \mathbf{Z}\mathcal{A} \rightarrow \mathcal{B}$ such that $T = \bar{T}J$. If \mathcal{A} is an additive category, there is a canonical projection $\theta: \mathbf{Z}\mathcal{A} \rightarrow \mathcal{A}$ such that $\theta J = \text{Id}$. If furthermore T is additive, then $\bar{T} = T\theta$.

Now suppose that a projective class \mathcal{P} in \mathcal{A} is given. For $A \in \mathcal{A}$, an augmented chain com-

plex $X \rightarrow A$ in $\mathbf{Z}\mathcal{A}$,

$$\rightarrow X_n \xrightarrow{d_n} X_{n-1} \rightarrow \dots \xrightarrow{d_1} X_0 \xrightarrow{d_0} A, \quad d_i d_{i+1} = 0,$$

is called a **\mathcal{P} -projective resolution** of A if (i) $X_n \in \mathcal{P}$ ($n \geq 0$), and (ii) it is **\mathcal{P} -acyclic** (i.e., the sequence

$$\begin{aligned} \rightarrow \mathbf{Z}\mathcal{A}(P, X_n) \xrightarrow{d_n} \mathbf{Z}\mathcal{A}(P, X_{n-1}) \rightarrow \dots \\ \xrightarrow{d_1} \mathbf{Z}\mathcal{A}(P, X_0) \xrightarrow{d_0} \mathbf{Z}\mathcal{A}(P, A) \rightarrow 0 \end{aligned}$$

is exact for any $P \in \mathcal{P}$ [16]). Note that the \mathcal{P} -acyclicity in this case implies the existence of a contracting homotopy

$$\begin{aligned} h_n: \mathbf{Z}\mathcal{A}(P, X_n) \rightarrow \mathbf{Z}\mathcal{A}(P, X_{n+1}), \\ n \geq -1, X_{-1} = A, \end{aligned}$$

satisfying $d_{n+1}h_n + h_{n-1}d_n = 1$.

Comparison theorem: Given two \mathcal{P} -projective resolutions $X \rightarrow A$ and $Y \rightarrow A$, we have that the chain complexes X , and Y , are chain equivalent in $\mathbf{Z}\mathcal{A}$.

Let $T: \mathcal{A} \rightarrow \mathcal{B}$ be a (covariant) functor with \mathcal{B} an Abelian category. The n th **left derived functor** $L_n T: \mathcal{A} \rightarrow \mathcal{B}$ ($n \geq 0$) of T , with respect to \mathcal{P} , is defined by $L_n T(A) = H_n(\bar{T}X)$ for a \mathcal{P} -projective resolution $X \rightarrow A$. The derived functors $L_n T$ will remain unaffected if we replace the projective class \mathcal{P} by its enlargement $\bar{\mathcal{P}} = \{\text{objects in } \mathcal{P} \text{ together with their retracts}\}$.

We can easily verify that (i) $L_0 T(P) = T(P)$ and $L_n T(P) = 0$ ($n > 0$) for $P \in \mathcal{P}$, and (ii) a short exact sequence $0 \rightarrow T' \rightarrow T \rightarrow T'' \rightarrow 0$ of functors $\mathcal{A} \rightarrow \mathcal{B}$ induces a long exact sequence of derived functors

$$\begin{aligned} \rightarrow L_n T' \rightarrow L_n T \rightarrow L_n T'' \xrightarrow{\partial} L_{n-1} T' \rightarrow \dots \\ \rightarrow L_0 T'' \rightarrow 0. \end{aligned}$$

If \mathcal{A} is preadditive and has a zero object and kernels, then it is routine to give a \mathcal{P} -projective resolution of any object. If \mathcal{A} has finite \dagger limits (i.e., \dagger finite products and \dagger equalizers), it can be proved that there exists a \mathcal{P} -projective resolution for any object [17].

There is a standard functorial construction which provides canonically a projective class \mathcal{P} in a category \mathcal{A} and a \mathcal{P} -projective resolution of any object in \mathcal{A} . Let (G, ε, δ) be a **cotriple** (or **comonad** [18] or functor coalgebra) in \mathcal{A} . Here $G: \mathcal{A} \rightarrow \mathcal{A}$ is an endofunctor, $\varepsilon: G \rightarrow \text{Id}$ and $\delta: G \rightarrow G^2 = GG$ are natural transformations such that $G\varepsilon \circ \delta = \varepsilon G \circ \delta = 1_G$ and $G\delta \circ \delta = \delta G \circ \delta$. A cotriple comes usually from a pair (F, U) of \dagger adjoint functors $U: \mathcal{A} \rightarrow \mathcal{C}$, $F: \mathcal{C} \rightarrow \mathcal{A}$ with natural bijection $\lambda: \mathcal{A}(F(C), A) \cong \mathcal{C}(C, U(A))$. Putting $G = FU: \mathcal{A} \rightarrow \mathcal{A}$, $\varepsilon(A) = \lambda^{-1}(1_{U(A)}): FU(A) \rightarrow A$, $\eta(C) = \lambda(1_{F(C)}): C \rightarrow UF(C)$, we have a cotriple $(G = FU, \varepsilon, \delta = F\eta U)$ in \mathcal{A} . Conversely, it is known that any cotriple in \mathcal{A} is induced from a suitable pair of adjoint functors.

Given a cotriple (G, ε, δ) in \mathcal{A} , we define a projective class $\mathcal{P}_G = \{G(A) \mid A \in \mathcal{A}\}$ (or its enlargement $\bar{\mathcal{P}}_G$) in \mathcal{A} , and an augmented simplicial object $G_*(A) \rightarrow A$ for any object A as follows. Put $G_n(A) = G^{n+1}(A)$ ($n \geq 0$), $\partial_i = \partial_i^n = G^i \varepsilon G^{n-i}(A): G_n(A) \rightarrow G_{n-1}(A)$ (face operator) and $\delta_i = \delta_i^n = G^i \delta G^{n-i}(A): G_n(A) \rightarrow G_{n+1}(A)$ (degeneracy operator) for $0 \leq i \leq n$, $G_{-1}(A) = A$. Then $G_*(A) \rightarrow A$ gives rise to a \mathcal{P}_G (or equivalently $\bar{\mathcal{P}}_G$)-projective resolution of A in $\mathbf{Z}\mathcal{A}$ with differentials $d_n = \sum_{i=0}^n (-1)^i \partial_i$. This is the bar resolution (or standard resolution) in a generalized sense, and there are many (co-)homology theories defined by means of such constructions.

Most of the above definitions and constructions can be dualized so as to give **injective classes**, **\mathcal{I} -injective resolutions**, and **right derived functors** with respect to \mathcal{I} , **triples** (or **manads**), etc. (See also [19] for generalized (co-)homology).

References

- [1] Sém. H. Cartan, 1950–1951, Paris, 1951.
- [2] H. P. Cartan and S. Eilenberg, *Homological algebra*, Princeton Univ. Press, 1956.
- [3] A. Grothendieck, Sur quelques points d'algèbre homologique, *Tôhoku Math. J.*, (2) 9 (1957), 119–221.
- [4] R. Godement, *Topologie algébrique et théorie des faisceaux*, Actualités Sci. Ind., Hermann, 1958.
- [5] D. G. Northcott, *An introduction to homological algebra*, Cambridge Univ. Press, 1960.
- [6] S. MacLane, *Homology*, Springer, 1963.
- [7] P. Freyd, *Abelian categories*, Harper, 1964.
- [8] A. Grothendieck (and J. Dieudonné), *Éléments de géométrie algébrique I–III*, Publ. Math. Inst. HES, 1960–1963.
- [9] J.-P. Serre, *Cohomologie galoisienne*, Lecture notes in math. 5, Springer, 1964.
- [10] S. Lang, *Rapport sur la cohomologie des groupes*, Benjamin, 1966.
- [11] E. Weiss, *Cohomology of groups*, Academic Press, 1969.
- [12] H. Bass, *Algebraic K-theory*, Benjamin, 1968.
- [13] D. Quillen, Projective modules over polynomial rings, *Inventiones Math.*, 36 (1976), 167–171.
- [14] S. Eilenberg and J. C. Moore, *Foundation of relative homological algebra*, Mem. Amer. Math. Soc., 55 (1965).
- [15] A. Dold and D. Puppe, *Homologie Nicht-Additiver Funktoren*, Anwendungen, Ann. Inst. Fourier, 11 (1961), 201–312.
- [16] M. Barr and J. Beck, *Homology and standard constructions*, Lecture notes in math. 80, Springer, 1969.

[17] M. Tierney and W. Vogel, *Simplicial derived functors*, Lecture notes in math. 86, Springer, 1969.

[18] S. MacLane, *Categories for the working mathematician*, Springer, 1971.

[19] J. F. Adams, *Stable homotopy and generalised homology*, Lecture notes, Univ. of Chicago, 1971.

201 (IX.6) Homology Theory

A. History

Homology theory is the oldest and most extensively developed portion of algebraic topology. Historically, it started with measuring the higher-dimensional **connectivity** of a space in the sense that the 0-dimensional connectivity is the number of connected components of the space. For example, take a 2-sphere S^2 and a 2-torus T^2 . Then T^2 is distinguished from S^2 by the fact that on T^2 a closed curve can be drawn without forming a boundary, while this is not true for S^2 . In fact, a curve (c_1 or c'_1 in Fig. 1) can be drawn on T^2 so that it does not form a boundary of an embedded 2-disk. On more complicated surfaces there are many kinds of such closed curves. The maximum number of such closed curves is the 1-dimensional connectivity of the surface; this is a topological property of the surface. For example, the 1-dimensional connectivity of S^2 is 0, that of T^2 is 2, and that of the surface in Fig. 2 is 6. A more general consideration of the bounding properties of q -dimensional closed submanifolds of a manifold led E. Betti (*Ann. Mat. Pure Appl.*, 4 (1871)) to introduce the notion of the q -dimensional connectivity of the manifold, which was a precursor of homology theory.

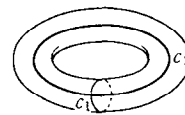


Fig. 1

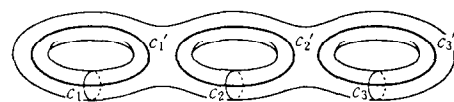


Fig. 2

The foundation of homology theory was laid by H. Poincaré [1]. He started his study of homology with analytic treatment of manifolds, which led to a series of complications.

Poincaré then introduced a new method for the study, now called *combinatorial topology*: He decomposed the manifold into elementary pieces or *cells*, which adjoin one another in a regular fashion; he then substituted the algebraic notions of *cycles* and *boundary operators* for the geometrical notions of closed submanifolds and boundaries. Thereby the notion of homology groups acquired an exact logical meaning, and the fundamental formulas, which are now called *Poincaré formulas*, were proved.

After Poincaré, much of the development of homology theory centered around the question of the **topological invariance of homology groups**, that is, the independence of the homology groups on the choice of *cellular decomposition*. Through the development of *simplicial complexes* and their techniques, J. W. Alexander (*Trans. Amer. Math. Soc.*, 28 (1926)) gave the first fully satisfactory proof for topological invariance of the homology groups of *polyhedra*. In those days, the homology groups themselves were barely recognized; instead, one dealt with numerical invariants such as the *Betti numbers* and the *torsion coefficients* [2, 3].

During the period 1925–1935 there was a gradual shift of interest from the numerical invariants to the homology groups themselves, and homology theory developed intensively [4, 5]. S. Lefschetz (*Trans. Amer. Math. Soc.*, 28 (1926)) added the theory of the *intersection products* to the homology of manifolds. He also invented *relative homology theory* (*Proc. Nat. Acad. Sci. US*, 13 (1927)) and generalized the *duality theorems* of Poincaré and Alexander (*ibid.*, 15 (1929)). G. de Rham (*J. Math. Pures Appl.*, 10 (1931)) obtained a duality theorem that relates the *exterior differential forms* in a manifold to the homology groups of the manifold. L. S. Pontryagin (*Ann. Math.*, (2) 35 (1934)) proved the complete group-invariant form of the Alexander duality theorem. These duality theorems seemed to reflect the existence of a theory dual to homology theory, and the genesis of this dual theory, now called *cohomology*, occurred in 1935, in the work of Alexander and A. N. Kolmogorov. It was discovered subsequently by Alexander (*Ann. Math.*, (2) 37 (1936)), E. Čech (*ibid.*), and H. Whitney (*ibid.*, 39 (1938)) that the cohomology of a polyhedron can be made into a ring.

On the other hand, after L. Vietoris (*Math. Ann.*, 97 (1927)) and P. S. Aleksandrov (*Ann. Math.*, (2) 30 (1928)), many devices were invented to extend the homology theory of polyhedra to general *topological spaces*, and numerous variants of homology theory appeared at the hands of Čech (1932), Lefschetz

(1933), Alexander (1935), and Kolmogorov (1936). This development served to clarify the relations between combinatorial and set-theoretic methods in topology (\rightarrow 426 Topology), whereas it produced complexity and confusion in homology theory [6]. S. Eilenberg and N. E. Steenrod (*Proc. Nat. Acad. Sci. US*, 31 (1945) and [7]) cleared the air by treating homology axiomatically.

Roughly speaking, a homology theory assigns *Abelian groups* to *topological spaces* and *homomorphisms* to *continuous mappings* of one space to another. In this way, a homology theory is an algebraic image of topology; it converts topological problems to algebraic problems. Starting from this viewpoint, Eilenberg and Steenrod selected some fundamental properties as axioms to characterize homology theory. This unified homology theories and allowed systematic treatment of homological problems which had previously been done separately “by hand” in each case. Moreover, it motivated the birth of a new branch of mathematics, called *homological algebra*.

B. Homology of Chain Complexes

A **chain complex** $C = \{C_q, \partial_q\}$ is a collection of (additive) Abelian groups C_q , one for each integer q , and of homomorphisms $\partial_q: C_q \rightarrow C_{q-1}$ such that $\partial_q \circ \partial_{q+1} = 0$ for each q . Elements of C_q are called *q-chains* of C , and ∂_q is called the **boundary operator**. A **subcomplex** $C' = \{C'_q, \partial'_q\}$ of C is a chain complex such that $C'_q \subset C_q$ and $\partial'_q = \partial_q|_{C'_q}$ for each q .

The *kernel* of ∂_q is denoted by $Z_q(C)$, and its element is called a *q-cycle* of C . The *image* of ∂_{q+1} is denoted by $B_q(C)$, and its element is called a *q-boundary* of C . The relation $\partial_q \circ \partial_{q+1} = 0$ implies $B_q(C) \subset Z_q(C)$. The *quotient group* $Z_q(C)/B_q(C)$ is denoted by $H_q(C)$, the *qth homology group* of C . Elements of $H_q(C)$ are called *q-dimensional homology classes* of C . Two cycles representing the same homology class are said to be **homologous**. The *direct sum* $\sum_q H_q(C)$ is denoted by $H_*(C)$ and is called the **homology group** of C .

If $C = \{C_q, \partial_q\}$ and $C' = \{C'_q, \partial'_q\}$ are chain complexes, a **chain mapping (chain map)** $\varphi: C \rightarrow C'$ is a sequence of homomorphisms $\varphi_q: C_q \rightarrow C'_q$ such that $\partial'_q \circ \varphi_q = \varphi_{q-1} \circ \partial_q$ for each q . If $\varphi: C \rightarrow C'$ is a chain mapping, then φ_q sends $Z_q(C)$ to $Z_q(C')$ and $B_q(C)$ to $B_q(C')$, and hence φ induces a homomorphism of $H_q(C)$ to $H_q(C')$.

If $H_q(C)$ is *finitely generated*, it can be decomposed into the direct sum of a free Abelian group $B_q(C)$ and a finite Abelian group $T_q(C)$. $B_q(C)$ and $T_q(C)$ are called the *qth Betti group*

of C and the q th **torsion group** of C , respectively. The rank ρ_q of $B_q(C)$ is called the q th **Betti number** of C . $T_q(C)$ is isomorphic to the direct sum of $\tau(q)$ finite cyclic groups of orders $\theta_1^q, \theta_2^q, \dots, \theta_{\tau(q)}^q$, where $\theta_i^q > 1$ and θ_i^q divides θ_{i+1}^q for $i = 1, \dots, \tau(q) - 1$ (\rightarrow 2 Abelian Groups). The numbers $\theta_1^q, \theta_2^q, \dots, \theta_{\tau(q)}^q$ are called the q th **torsion coefficients** of C .

If $H_*(C)$ is finitely generated, then the number $\chi(C) = \sum_q (-1)^q \rho_q$ is called the **Euler number**, the **Euler characteristic**, or the **Euler-Poincaré characteristic** of C . In this case a \dagger polynomial $\sum_q \rho_q t^q$ with variable t is called the **Poincaré polynomial** of C .

Let C be a chain complex such that, for each q , C_q is a \dagger free Abelian group of finite \dagger rank. Then the q th Betti number and the q th torsion coefficients are well defined for each q . Denote the ranks of C_q and $B_q(C)$ by α_q and β_q , respectively. Then it holds that $\rho_q = \alpha_q - \beta_q - \beta_{q-1}$, and hence $\chi(C) = \sum_q (-1)^q \alpha_q$. (**Euler-Poincaré formula**). Moreover, there exists a set of bases, one for each C_q , with the following properties: For each q , the base for C_q is composed of five types of elements, a_i^q ($1 \leq i \leq \beta_q - \tau_q$), b_i^q ($1 \leq i \leq \tau_q$), c_i^q ($1 \leq i \leq \rho_q$), d_i^q ($1 \leq i \leq \tau_{q-1}$), and e_i^q ($1 \leq i \leq \beta_{q-1} - \tau_{q-1}$); ∂_q satisfies $\partial a_i^q = 0$, $\partial b_i^q = 0$, $\partial c_i^q = 0$, $\partial d_i^q = \theta_i^{q-1} b_i^{q-1}$, and $\partial e_i^q = a_i^{q-1}$. Such a set of bases is called the **canonical basis** of C .

Let C be a chain complex such that each C_q is a free Abelian group with a given base $\{\sigma_i^q\}$. Then the **incidence number** $[\sigma_i^q : \sigma_j^{q-1}] \in \mathbf{Z}$ is defined by $\partial_q(\sigma_i^q) = \sum_j [\sigma_i^q : \sigma_j^{q-1}] \sigma_j^{q-1}$. This notion was commonly used in the early days of topology.

C. Homology of Simplicial Complexes

Let K be a \dagger simplicial complex. An **oriented q -simplex** σ of K is a q -simplex $s \in K$ together with an equivalence class of \dagger total orderings of the vertices of s , two orderings being equivalent if they differ by an even permutation of the vertices. If a_0, \dots, a_q are the vertices of s , then $[a_0, a_1, \dots, a_q]$ denotes the oriented q -simplex of K consisting of the simplex s together with the equivalence class of the ordering $a_0 < a_1 < \dots < a_q$ of its vertices. For every vertex a of K there is a unique oriented 0-simplex $[a]$, and to every q -simplex with $q \geq 1$ there correspond exactly two oriented q -simplexes, which are said to be **opposites** of one another.

Let $C_q(K)$ denote the Abelian group generated by the oriented q -simplexes of K with the relations $\sigma + \sigma' = 0$ if σ' is the opposite of σ . If we choose an oriented q -simplex σ_i^q for each q -simplex s_i^q of K , then each element of $C_q(K)$ is written uniquely as a finite sum $\sum_i g_i \sigma_i^q$ with integers $g_i \neq 0$, and

$C_q(K)$ is a free Abelian group generated by the set $\{\sigma_i^q\}$. We define a homomorphism $\partial_q : C_q(K) \rightarrow C_{q-1}(K)$ by $\partial_q[a_0, a_1, \dots, a_q] = \sum_{i=0}^q (-1)^i [a_0, \dots, a_{i-1}, a_{i+1}, \dots, a_q]$. Then $\partial_q \circ \partial_{q+1} = 0$ holds, and we have a chain complex $C(K) = \{C_q(K), \partial_q\}$ ($C_q(K) = \{0\}$ if $q < 0$), called the **(oriented) simplicial chain complex**. The homology group $H_*(C(K))$ is denoted by $H_*(K)$ and is called the **(integral) homology group of the simplicial complex K** .

Let K_1 and K_2 be simplicial complexes, and let $f : K_1 \rightarrow K_2$ be a \dagger simplicial mapping. Then for each q , a homomorphism $f_{\#q} : C_q(K_1) \rightarrow C_q(K_2)$ can be defined by $f_{\#q}([a_0, a_1, \dots, a_q]) = [f(a_0), f(a_1), \dots, f(a_q)]$, where the right-hand side is understood to be 0 if $f(a_0), f(a_1), \dots, f(a_q)$ are not distinct. The sequence $f_{\#} = \{f_{\#q}\}$ is a chain mapping of $C(K_1)$ to $C(K_2)$, and it induces a homomorphism $f_* : H_*(K_1) \rightarrow H_*(K_2)$.

If K is a \dagger finite simplicial complex, then the Betti numbers, the torsion coefficients, and the Euler characteristic of K are defined to be those for the chain complex $C(K)$.

Let K_1 and K_2 be subcomplexes of a simplicial complex K . Then we have the following \dagger exact sequence which relates the homology group of $K_1 \cup K_2$ to the homology groups of K_1 , K_2 , and $K_1 \cap K_2$: $\dots \xrightarrow{\alpha} H_q(K_1 \cap K_2) \xrightarrow{\beta} H_q(K_1) + H_q(K_2) \xrightarrow{\gamma} H_q(K_1 \cup K_2) \xrightarrow{\delta} H_{q-1}(K_1 \cap K_2) \xrightarrow{\alpha} \dots$, where α , β , and δ_* are defined as follows. Let $i_l : K_l \cap K_2 \rightarrow K_1$ and $j_l : K_l \rightarrow K_1 \cup K_2$ ($l = 1, 2$) be inclusion mappings; then $\alpha(a) = (i_{1*}(a), -i_{2*}(a))$ and $\beta(a_1, a_2) = j_{1*}(a_1) + j_{2*}(a_2)$. If $z = c_1 + c_2$ ($c_i \in C(K_i)$) is a cycle of $C(K_1 \cup K_2)$ then $\partial(c_1) = -\partial(c_2)$ is a cycle of $C(K_1 \cap K_2)$; δ_* sends the homology class of z to the homology class of $\partial(c_1)$. The sequence is called the **Mayer-Vietoris exact sequence** of the couple $\{K_1, K_2\}$, and δ_* is referred to as the **connecting homomorphism**. The prototype of the Mayer-Vietoris exact sequence was obtained by W. Mayer (*Monatsh. Math. Phys.*, 36 (1929)) and L. Vietoris (*ibid.*, 37 (1930)). The present form is due to Eilenberg and Steenrod [7].

D. Homology of Polyhedra

If K is a \dagger simplicial complex and K' is a \dagger subdivision of K , then there exists a canonical isomorphism $H_*(K) \cong H_*(K')$. This proves that if K_1 and K_2 are \dagger simplicial decompositions of a \dagger polyhedron then $H_*(K_1)$ and $H_*(K_2)$ are isomorphic, because there exists a common subdivision of K_1 and K_2 . Thus we may define the **(integral) homology groups $H_*(X)$ of a polyhedron X** to be the homology group $H_*(K)$ of a simplicial decomposition K of X .

Let X and Y be polyhedra, and let $f : X \rightarrow Y$ be a continuous mapping. Take a \dagger simplicial approximation $\varphi : K \rightarrow L$ of f . Then a homo-

morphism of $H_*(X)$ to $H_*(Y)$ given by the induced homomorphism $\varphi_*: H_*(K) \rightarrow H_*(L)$ is independent of the choice of φ , and is denoted by f_* . The following properties hold: (i) $1_* = 1: H_*(X) \rightarrow H_*(X)$, where 1 is the identity; (ii) If $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ are continuous mappings, then $(g \circ f)_* = g_* \circ f_*: H_*(X) \rightarrow H_*(Z)$; (iii) If $f, f': X \rightarrow Y$ are homotopic, then $f_* = f'_*: H_*(X) \rightarrow H_*(Y)$. These imply the **homotopy invariance of the homology group** stated as follows: If X and Y are polyhedra which are \dagger homotopy equivalent, then $H_*(X)$ and $H_*(Y)$ are isomorphic. Specifically, the homology group is a topological invariant. Thus if X is a \dagger triangulable space (for example, if X is a \dagger differentially manifold), then its homology group $H_*(X)$ can be defined to be the homology group $H_*(K)$, where (K, t) is a \dagger triangulation of X . This homology group is referred to as the **simplicial homology group** of X . Similarly, if X is a compact triangulable space, the Betti numbers $\rho_q(X)$ of X , etc. can be defined to be those for the chain complex $C(K)$.

If we denote by pt a single point, then $H_0(pt) = \mathbb{Z}$ (the group of integers) and $H_q(pt) = 0$ if $q \neq 0$. For an n -sphere S^n , a 2-torus T^2 , and a \dagger real projective plane P^2 , the homology groups can be computed as follows by means of their triangulations: (1) $H_0(S^n) \cong H_n(S^n) \cong \mathbb{Z}$, and $H_q(S^n) = 0$ if $q \neq 0, n$; (2) $H_0(T^2) \cong H_2(T^2) \cong \mathbb{Z}$, $H_1(T^2) \cong \mathbb{Z} + \mathbb{Z}$, and $H_q(T^2) = 0$ if $q \neq 0, 1, 2$; (3) $H_0(P^2) \cong \mathbb{Z}$, $H_1(P^2) \cong \mathbb{Z}_2$, and $H_q(P^2) = 0$ if $q \neq 0, 1$, where $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$.

Two surfaces are homeomorphic if and only if their integral homology groups are isomorphic (\rightarrow 410 Surfaces).

E. Singular Homology

There are various devices for defining homology groups of general topological spaces. A familiar one is the **singular homology theory** initiated by S. Lefschetz (*Bull. Amer. Math. Soc.*, 39 (1933)) and improved by S. Eilenberg (*Ann. Math.*, (2) 45 (1944)).

The **standard q -simplex** is the convex set $\Delta^q \subset \mathbb{R}^{q+1}$ consisting of all $(q+1)$ -tuples (t_0, t_1, \dots, t_q) of real numbers with $t_i \geq 0$, $t_0 + t_1 + \dots + t_q = 1$. Any continuous mapping of Δ^q to a topological space X is called a **singular q -simplex** in X . The i th **face** of a singular q -simplex $\sigma: \Delta^q \rightarrow X$ is the singular $(q-1)$ -simplex $\sigma \circ \varepsilon_i: \Delta^{q-1} \rightarrow X$, where the linear embedding $\varepsilon_i: \Delta^{q-1} \rightarrow \Delta^q$ is defined by $\varepsilon_i(t_0, \dots, t_{i-1}, t_{i+1}, \dots, t_q) = (t_0, \dots, t_{i-1}, 0, t_{i+1}, \dots, t_q)$.

For each integer q , let $S_q(X)$ denote the free Abelian group generated by the singular q -simplexes in X ($S_q(X) = 0$ if $q < 0$), and define a homomorphism $\partial_q: S_q(X) \rightarrow S_{q-1}(X)$ by $\partial_q(\sigma) =$

$\sum_{i=0}^q (-1)^i \sigma \circ \varepsilon_i$. Then we have a chain complex $S(X) = \{S_q(X), \partial_q\}$, called the **singular chain complex** of X . The homology group $H_*(S(X))$ is denoted by $H_*(X)$ and is called the **integral singular homology group** of the topological space X .

Given a continuous mapping $f: X \rightarrow Y$, a chain mapping $f_*: S(X) \rightarrow S(Y)$ is defined by sending each singular simplex $\sigma: \Delta^q \rightarrow X$ to the singular simplex $f \circ \sigma: \Delta^q \rightarrow Y$, and it induces the homomorphism $f_*: H_*(X) \rightarrow H_*(Y)$. The properties (i), (ii), (iii) of f_* in Section D hold for continuous mappings of topological spaces, and hence the singular homology group is a homotopy invariant.

The homology group $H_*(K)$ of a simplicial complex K is isomorphic to the singular homology group $H_*(|K|)$ of the polyhedron $|K|$. Therefore the simplicial homology group of a triangulable space is isomorphic to the singular homology group of the space.

If $\{X_i\}$ is the set of \dagger arcwise connected components of a topological space X , then $H_*(X) \cong \sum_i H_*(X_i)$. If X is arcwise connected, then $H_0(X) \cong \mathbb{Z}$. If $\{A_\lambda\}$ is the collection of all the compact subsets of X directed by inclusion, then $H_*(X)$ is isomorphic to the \dagger inductive limit $\varinjlim H_*(A_\lambda)$. It is not true that there is a Mayer-Vietoris sequence in singular homology for any couple $\{X_1, X_2\}$ of subsets of X . However, for certain couples $\{X_1, X_2\}$, there is a Mayer-Vietoris exact sequence of $\{X_1, X_2\}$: $\dots \xrightarrow{\partial_q} H_q(X_1 \cap X_2) \xrightarrow{\alpha} H_q(X_1) + H_q(X_2) \xrightarrow{\beta} H_q(X_1 \cup X_2) \xrightarrow{\partial_{q-1}^*} H_{q-1}(X_1 \cap X_2) \xrightarrow{\alpha} \dots$. For example, this holds if $X = \text{Int } X_1 \cup \text{Int } X_2$, where Int denotes the \dagger interior.

Let $c: X \rightarrow pt$ be the mapping of a topological space to a single point. Then the kernel of $c_*: H_*(X) \rightarrow H_*(pt)$ is denoted by $\tilde{H}_*(X)$ and is called the **reduced homology group** of X . It holds that $H_*(X) \cong \tilde{H}_*(X) + H_*(pt)$. Regard the \dagger suspension SX as the union of two copies of the \dagger cone over X . Then the connecting homomorphism in the Meyer-Vietoris sequence gives an isomorphism $\tilde{H}_q(SX) \cong \tilde{H}_{q-1}(X)$ for any q . The inverse of this isomorphism is called the **suspension isomorphism for homology**.

Let M be a $\dagger C^\infty$ -manifold. A C^∞ -singular q -simplex in M is a singular q -simplex $\sigma: \Delta^q \rightarrow M$ such that σ extends to a C^∞ -mapping from an open neighborhood of Δ^q in $\{(t_0, t_1, \dots, t_q) \in \mathbb{R}^{q+1} \mid t_0 + t_1 + \dots + t_q = 1\}$ to M . The totality of C^∞ -singular simplexes in M generates a subcomplex of $S(M)$, denoted by $S^\infty(M)$. The inclusion $S^\infty(M) \subset S(M)$ induces an isomorphism $H_*(S^\infty(M)) \cong H_*(M)$.

Let M be an n -dimensional \dagger topological manifold. Then $H_q(M) = 0$ unless $0 \leq q \leq n$, and $H_*(M)$ is finitely generated if M is compact.

F. Homology of CW Complexes

Homology theory is tractable in the category of \dagger CW complexes by virtue of the facts stated below.

Let X be a topological space and A its subset. Then we denote by X/A the quotient space obtained from X by shrinking A to a point, understanding X/\emptyset to be the disjoint union $X \cup pt$. If X_1, X_2 are \dagger subcomplexes of a CW complex, then the Mayer-Vietoris exact sequence of $\{X_1, X_2\}$ and the **excision isomorphism** $i_*: H_*(X_1/(X_1 \cap X_2)) \cong H_*((X_1 \cup X_2)/X_2)$ (i : inclusion) are valid. If A is a subcomplex of a CW complex X , then we have the following **reduced homology exact sequence** of $(X, A): \dots \xrightarrow{\partial_*} \tilde{H}_q(X) \xrightarrow{i_*} \tilde{H}_q(X/A) \xrightarrow{j_*} \tilde{H}_{q-1}(A) \xrightarrow{i_*} \dots$, where i_* and j_* are induced by the inclusion $i: A \rightarrow X$ and the collapsing $j: X \rightarrow X/A$, and ∂_* is given by a commutative diagram

$$\begin{array}{ccccc} \tilde{H}_q(X/A) & \xrightarrow{\partial_*} & \tilde{H}_{q-1}(A) \\ \cong \uparrow h_* & & \cong \downarrow S \\ \tilde{H}_q(X \cup CA) & \xrightarrow{h_*} & \tilde{H}_q(SA) \end{array}$$

Here CA is the cone over A , S is the suspension isomorphism, and $h: X \cup CA \rightarrow (X \cup CA)/CA = X/A$ and $h': X \cup CA \rightarrow (X \cup CA)/X = SA$ are collapsings. More generally, if A, B are subcomplexes of a CW complex X and $A \supset B$, then we have the following **reduced homology exact sequence** of $(X, A, B): \dots \xrightarrow{\partial_*} \tilde{H}_q(A/B) \xrightarrow{i_*} \tilde{H}_q(X/B) \xrightarrow{j_*} \tilde{H}_q(X/A) \xrightarrow{k_*} \tilde{H}_{q-1}(A/B) \xrightarrow{i_*} \dots$. Furthermore, the homology group $H_*(X)$ of a CW complex X can be computed in the following manner.

Let X^q denote the q -skeleton of X , i.e., the union of all cells of dimensions $\leq q$. Put $C_q(X) = \tilde{H}_q(X^q/X^{q-1})$, and let $\partial_q: C_q(X) \rightarrow C_{q-1}(X)$ be the connecting homomorphism $\partial_*: \tilde{H}_q(X^q/X^{q-1}) \rightarrow \tilde{H}_{q-1}(X^{q-1}/X^{q-2})$ in the reduced homology exact sequence of (X^q, X^{q-1}, X^{q-2}) . Then $C(X) = \{C_q(X), \partial_q\}$ is a chain complex such that $C_q(X)$ is a free Abelian group with one generator for each q -cell of X . If X is a polyhedron $|K|$, then $C(X)$ coincides with the simplicial chain complex $C(K)$. The homology group $H_*(C(K))$ is called the **cellular homology group** of the CW complex X . This is isomorphic to the singular homology group $H_*(X)$.

Since a CW decomposition frequently requires fewer cells than a simplicial decomposition, the cellular homology groups are useful in calculating the homology groups. For example, the \dagger complex n -dimensional projective space CP^n has a CW decomposition with a single $2i$ -cell for each $i = 0, 1, \dots, n$, and hence we see immediately that $H_q(CP^n) \cong \mathbb{Z}$ if $q = 2i$ ($0 \leq i \leq n$) and $= 0$ otherwise.

If X is a finite CW complex and α_q denotes

the number of q -cells of X , then we have the **Euler-Poincaré formula** $\chi(X) = \sum_q (-1)^q \alpha_q$. In particular, if X is homeomorphic to S^2 then we have the **Euler theorem on polyhedra**: $\alpha_0 - \alpha_1 + \alpha_2 = 2$. This was the first important result in topology (L. Euler, 1752).

G. Homology with Coefficients in Abelian Groups

Given a chain complex C and an Abelian group G , we have a new chain complex $C \otimes G$ given by $(C \otimes G)_q = C_q \otimes G$ and $\partial_q(c \otimes g) = \partial_q c \otimes g$ ($c \in C_q, g \in G$), where \otimes is the \dagger tensor product of Abelian groups. For a topological space X , the homology group of the chain complex $S(X) \otimes G$ is denoted by $H_*(X; G)$ and is called the **singular homology group of X with coefficients in G** . The **homology group $H_*(K; G)$ of a simplicial complex K with coefficients in G** and the **cellular homology group $H_*(C(X); G)$ of a CW complex X with coefficients in G** are similarly defined. The homology group with coefficients in \mathbb{Z} is the integral homology group.

The previous results for the integral homology groups generalize in a straightforward fashion to homology groups with coefficients in G .

We have the homomorphism $\kappa: H_q(X) \otimes G \rightarrow H_q(X; G)$ sending $a \otimes g \in H_q(X) \otimes G$ to the homology class of $z \otimes g \in Z_q(S(X) \otimes G)$, where z is a representative cycle of a . The following theorem is known as the **universal coefficient theorem for homology**, since it expresses $H_q(X; G)$ in terms of $H_q(X)$, $H_{q-1}(X)$, and G : There is an exact sequence $0 \rightarrow H_q(X) \otimes G \xrightarrow{\kappa} H_q(X; G) \rightarrow \text{Tor}(H_{q-1}(X), G) \rightarrow 0$, and this sequence is split (\rightarrow 200 Homological Algebra). Universal coefficient theorems of this type were first shown by S. Eilenberg and S. MacLane (*Ann. Math.*, (2) 43 (1942)).

Let Λ be a \dagger ring with a unit 1. Then a **chain complex over Λ** is a chain complex C such that each C_q is a $\dagger\Lambda$ -module and each ∂_q is a $\dagger\Lambda$ -homomorphism. The homology groups $H_q(C)$ of a chain complex C over Λ are Λ -modules. If C is a chain complex, $C \otimes \Lambda$ forms naturally a chain complex over Λ . In particular, if X is a topological space, then $S(X) \otimes \Lambda$ is a chain complex over Λ , and $H_q(X; \Lambda)$ are Λ -modules. In this case, the induced homomorphisms $f_*: H_q(X; \Lambda) \rightarrow H_q(Y; \Lambda)$ are Λ -homomorphisms. The homology groups with coefficients in a field \mathbf{k} are \dagger vector spaces over \mathbf{k} and are useful in applications. If $H_*(X)$ is finitely generated, then $\chi(X) = \sum_q (-1)^q \dim_{\mathbf{k}} H_q(X; \mathbf{k})$ holds for any field \mathbf{k} .

H. Cohomology

A **cochain complex** $C = \{C^q, \delta^q\}$ is a collection of Abelian groups C^q , one for each integer q , and of homomorphisms $\delta^q: C^q \rightarrow C^{q+1}$ such that $\delta^{q+1} \circ \delta^q = 0$. Elements of C^q are called q -**cochains**, and δ^q is called the **coboundary operator**. The notions of **subcomplex of a cochain complex**, **cocycle**, **coboundary**, **cohomology group**, and **cochain mapping** are defined as in chain complex. Let $\text{Hom}(A, B)$ denote the \dagger group of homomorphisms from an Abelian group A to an Abelian group B . Given a chain complex C and an Abelian group G , a cochain complex $C^* = \text{Hom}(C, G)$ is defined by $C^q = \text{Hom}(C_q, G)$ and $(\delta^q u)(c) = u(\partial_{q+1} c)$ ($u \in C^q, c \in C_{q+1}$).

For a topological space X , the cochain complex $\text{Hom}(S(X), G)$ is called the **singular cochain complex** of X with coefficients in G , and elements of $\text{Hom}(S_q(X), G)$ are called **singular q -cochains** of X . The cohomology group of $\text{Hom}(S(X), G)$ is denoted by $H^*(X; G)$ and is called the **singular cohomology group of X with coefficients in G** . We write $H^*(X)$ for $H^*(X; \mathbf{Z})$; this is called the **integral cohomology group of X** . Similarly, the **cohomology group $H^*(K; G)$ of a simplicial complex K with coefficients in G** and the **cellular cohomology group $H^*(C(X); G)$ of a CW complex X with coefficients in G** are defined. There are isomorphisms $H^*(K; G) \cong H^*(|K|; G)$ and $H^*(C(X); G) \cong H^*(X; G)$.

If $f: X \rightarrow Y$ is a continuous mapping, then a cochain map $f^*: \text{Hom}(S(Y), G) \rightarrow \text{Hom}(S(X), G)$ is defined by $(f^* u)(c) = u(f_\# c)$ with $u \in \text{Hom}(S_q(Y), G)$ and $c \in S_q(X)$. Therefore f induces the homomorphism $f^*: H^*(Y; G) \rightarrow H^*(X; G)$. The following properties hold: (i) $1^* = 1$; (ii) $(g \circ f)^* = f^* \circ g^*$; (iii) If f and f' are homotopic, then $f^* = f'^*$. In particular, the singular cohomology groups are homotopy invariants. The 'cokernel of $c^*: H^*(pt; G) \rightarrow H^*(X; G)$ induced by the mapping $c: X \rightarrow pt$ is denoted by $\tilde{H}^*(X; G)$ and is called the **reduced cohomology group of X** .

For $\xi \in H^q(X; G)$ and $a \in H_q(X)$, the **Kronecker index** $\langle \xi, a \rangle \in G$ is defined naturally in terms of representatives of ξ and a . We have the following **universal coefficient theorem for cohomology**: There is an exact sequence $0 \rightarrow \text{Ext}(H_{q-1}(X), G) \rightarrow H^q(X; G) \xrightarrow{\kappa} \text{Hom}(H_q(X), G) \rightarrow 0$, and this sequence is split, where κ is given by the Kronecker index (\rightarrow 200 Homological Algebra).

If Λ is a ring with 1, then a **cochain complex over Λ** is defined analogously to a chain complex over Λ . The singular cochain complex $\text{Hom}(S(X), \Lambda)$ forms naturally a cochain complex over Λ , and $H^q(X; \Lambda)$ are Λ -modules. For $\xi \in H^q(X; \Lambda)$ and $a \in H_q(X; \Lambda)$, the **Kronecker**

index $\langle \xi, a \rangle \in \Lambda$ is defined naturally. If \mathbf{k} is a field, then the vector spaces $H^q(X; \mathbf{k})$ are identified with the dual space of $H_q(X, \mathbf{k})$ by means of the Kronecker index.

If M is a C^∞ -manifold, then there is a cochain complex $\mathfrak{D}(M) = \{\mathfrak{D}^q(M), d\}$ over the field \mathbf{R} of real numbers, where $\mathfrak{D}^q(M)$ is the real vector space consisting of the \dagger differential forms of degree q on M , and $d: \mathfrak{D}^q(M) \rightarrow \mathfrak{D}^{q+1}(M)$ is the \dagger exterior differentiation. The cochain complex $\mathfrak{D}(M)$ is called the **de Rham complex** of M , and its cohomology group $H^*(\mathfrak{D}(M))$ is called the **de Rham cohomology group of M** . A cochain mapping $\mathcal{J}: \mathfrak{D}(M) \rightarrow \text{Hom}(S^\infty(M), \mathbf{R})$ is defined by

$$(\mathcal{J}(\omega))(\sigma) = \int_{\Delta^q} \sigma^* \omega,$$

where $\omega \in \mathfrak{D}^q(M)$, $\sigma: \Delta^q \rightarrow M$ is a C^∞ singular q -simplex in M , and $\sigma^* \omega$ denotes the \dagger pullback of ω by σ . We have isomorphisms

$$H^*(\mathfrak{D}(M)) \xrightarrow{\mathcal{J}} H^*(\text{Hom}(S^\infty(M), \mathbf{R})) \xleftarrow{i^*} H^*(M; \mathbf{R}),$$

where i^* is induced by the inclusion $S^\infty(M) \subset S(M)$. This result is called the **de Rham theorem** on the cohomology of manifolds (\rightarrow 105 Differentiable Manifolds).

I. Cohomology Rings

Given a topological space X and a ring Λ , the **cup product** $u \smile v \in \text{Hom}(S_{p+q}(X), \Lambda)$ of cochains $u \in \text{Hom}(S_p(X), \Lambda)$ and $v \in \text{Hom}(S_q(X), \Lambda)$ is defined by $(u \smile v)(\sigma) = u(\sigma \circ \varepsilon)v(\sigma \circ \varepsilon')$, where $\sigma: \Delta^{p+q} \rightarrow X$ is a singular $(p+q)$ -simplex, $\varepsilon: \Delta^p \rightarrow \Delta^{p+q}$ and $\varepsilon': \Delta^q \rightarrow \Delta^{p+q}$ are given by $\varepsilon(t_0, t_1, \dots, t_p) = (t_0, t_1, \dots, t_p, 0, \dots, 0)$, and $\varepsilon'(t_p, t_{p+1}, \dots, t_{p+q}) = (0, \dots, 0, t_p, t_{p+1}, \dots, t_{p+q})$. The product operation is bilinear, and the formula $\delta(u \smile v) = \delta u \smile v + (-1)^p u \smile \delta v$ holds. Therefore it gives rise to the **cup product** $\xi \smile \eta \in H^{p+q}(X; \Lambda)$ of cohomology classes $\xi \in H^p(X; \Lambda)$ and $\eta \in H^q(X; \Lambda)$. This cup product operation makes $H^*(X; \Lambda)$ into a ring, which is called the **singular cohomology ring** of X with coefficients in Λ . If Λ has 1, the cohomology class represented by the 0-cocycle taking the value 1 on each singular 0-simplex serves as 1 of $H^*(X; \Lambda)$. If Λ is commutative, then $\xi \eta = (-1)^{pq} \eta \xi$ holds. The induced homomorphism $f^*: H^*(Y; \Lambda) \rightarrow H^*(X; \Lambda)$ preserves the product, and hence the cohomology ring is a homotopy invariant.

If K is a simplicial complex, the **cup product** operation $\smile: H^p(K; \Lambda) \otimes H^q(K; \Lambda) \rightarrow H^{p+q}(K; \Lambda)$ is induced from the operation $\smile: \text{Hom}(C_p(K), \Lambda) \otimes \text{Hom}(C_q(K), \Lambda) \rightarrow \text{Hom}(C_{p+q}(K), \Lambda)$ defined as follows. Adopting a \dagger linear ordering of vertices of K , we write all oriented simplexes in this ordering. Then,

for $u \in \text{Hom}(C_p(K), \Lambda)$ and $v \in \text{Hom}(C_q(K), \Lambda)$, we define $u \smile v \in \text{Hom}(C_{p+q}(K), \Lambda)$ by $(u \smile v)([a_0, a_1, \dots, a_{p+q}]) = u([a_0, a_1, \dots, a_p])v([a_{p+1}, \dots, a_{p+q}])$. The canonical isomorphism from the cohomology of K to the singular cohomology of $|K|$ preserves the cup product.

On the de Rham complex $\mathfrak{D}(M)$ of a C^∞ -manifold M , we have the exterior product $\omega \wedge \theta \in \mathfrak{D}^{p+q}(M)$ of $\omega \in \mathfrak{D}^p(M)$ and $\theta \in \mathfrak{D}^q(M)$. This makes $H^*(\mathfrak{D}(M))$ into a ring, which is called the **de Rham cohomology ring**. The canonical isomorphism of $H^*(\mathfrak{D}(M))$ to $H^*(M; \mathbb{R})$ preserves the product (\rightarrow 105 Differentiable Manifolds).

Examples. (1) Let $T^n = S^1 \times \dots \times S^1$ denote the n -dimensional torus, and let $\pi_i: T^n \rightarrow S^1$ denote the projection to the i th factor ($1 \leq i \leq n$). Take a generator ξ of the Λ -module $H^1(S^1; \Lambda)$, and put $\xi_i = \pi_i^*(\xi) \in H^1(T^n; \Lambda)$. Then $H^*(T^n; \Lambda)$ is the exterior algebra over Λ generated by ξ_1, \dots, ξ_n . (2) If we denote by CP^n the complex n -dimensional projective space, then $H^q(CP^n; \Lambda)$ is Λ if $q = 2i$ ($0 \leq i \leq n$) and 0 otherwise. If ξ is a generator of the Λ -module $H^2(CP^n; \Lambda)$, then ξ^i generates the Λ -module $H^{2i}(CP^n; \Lambda)$ ($0 \leq i \leq n$). Thus $H^*(CP^n; \Lambda)$ is the quotient ring $\Lambda[\xi]/(\xi^{n+1})$ of the polynomial ring $\Lambda[\xi]$ by the ideal (ξ^{n+1}) . (3) If P^n denote the real n -dimensional projective space, then $H^*(P^n; \mathbb{Z}_2) \cong \mathbb{Z}_2[\xi]/(\xi^{n+1})$, where ξ is the generator of $H^1(P^n; \mathbb{Z}_2)$.

J. Homology of Product Spaces

If C and D are chain complexes, their **tensor product** $C \otimes D$ is a chain complex given by $(C \otimes D)_n = \sum_{p+q=n} C_p \otimes D_q$, and $\partial_n(c \otimes d) = \partial_p(c) \otimes d + (-1)^p c \otimes \partial_q(d)$ ($c \in C_p, d \in D_q$). The following **Eilenberg-Zilber theorem** (*Amer. J. Math.*, 75 (1953)) is the link between the algebra of tensor products and the geometry of product spaces: For the product space $X \times Y$ of topological spaces X and Y , there is an isomorphism $\rho_*: H_*(X \times Y; \mathbb{G}) \cong H_*(S(X) \otimes S(Y) \otimes \mathbb{G})$ induced from a chain mapping $\rho: S(X \times Y) \rightarrow S(X) \otimes S(Y)$ defined as follows: Given a singular n -simplex $\sigma: \Delta^n \rightarrow X \times Y$, we define for each p ($0 \leq p \leq n$) a singular p -simplex σ'_p in X to be the composite $\Delta^p \xrightarrow{\pi_1} \Delta^n \xrightarrow{\sigma} X \times Y \xrightarrow{\pi_1} X$, where $\varepsilon(t_0, t_1, \dots, t_p) = (t_0, t_1, \dots, t_p, 0, \dots, 0)$ and π_1 is the projection to the first factor. Similarly, we define for each q ($0 \leq q \leq n$) a singular q -simplex σ''_q in Y to be the composite $\Delta^q \xrightarrow{\pi_2} \Delta^n \xrightarrow{\sigma} X \times Y \xrightarrow{\pi_2} Y$, where $\varepsilon'(t_{n-q}, \dots, t_n) = (0, \dots, 0, t_{n-q}, \dots, t_n)$ and π_2 is the projection to the second factor. Then ρ is defined by $\rho(\sigma) = \sum_{p+q=n} \sigma'_p \otimes \sigma''_q$ and is called the **Alexander-Whitney mapping** (**Alexander-Whitney map**).

Given a ring Λ , a chain mapping $\mu: (S(X) \otimes$

$S(Y) \otimes \Lambda) \rightarrow S(X) \otimes S(Y) \otimes \Lambda$ is defined by $\mu((c \otimes \lambda') \otimes (d \otimes \lambda'')) = c \otimes d \otimes \lambda' \lambda''$; it induces homomorphisms $\mu_*: H_p(X; \Lambda) \otimes H_q(Y; \Lambda) \rightarrow H_{p+q}(S(X) \otimes S(Y) \otimes \Lambda)$. The **cross product** $a \times b \in H_{p+q}(X \times Y; \Lambda)$ of $a \in H_p(X; \Lambda)$ and $b \in H_q(Y; \Lambda)$ is defined to be $\rho_*^{-1}(\mu_*(a \otimes b))$. If Λ is a commutative ring with 1, then the cross product defines a Λ -homomorphism $\times: H_p(X; \Lambda) \otimes_\Lambda H_q(Y; \Lambda) \rightarrow H_{p+q}(X \times Y; \Lambda)$, and satisfies $(a \times b) \times c = a \times (b \times c)$, $T_*(a \times b) = (-1)^{pq} b \times a$, $(f \times g)_*(a \times b) = f_*(a) \times g_*(b)$, where $T: X \times Y \rightarrow Y \times X$ is the mapping interchanging factors, and $f: X \rightarrow X'$, $g: Y \rightarrow Y'$ are continuous mappings. If Λ is a principal ideal domain, there is an exact sequence

$$0 \rightarrow \sum_{p+q=n} H_p(X; \Lambda) \otimes_\Lambda H_q(Y; \Lambda) \xrightarrow{\times} H_n(X \times Y; \Lambda) \rightarrow \sum_{p+q=n-1} \text{Tor}_\Lambda(H_p(X; \Lambda), H_q(Y; \Lambda)) \rightarrow 0,$$

and this sequence is split (\rightarrow 200 Homological Algebra). In particular, if \mathbf{k} is a field we have the following isomorphism of vector spaces:

$$\times: \sum_{p+q=n} H_p(X; \mathbf{k}) \otimes_{\mathbf{k}} H_q(Y; \mathbf{k}) \cong H_n(X \times Y; \mathbf{k}).$$

This is called the **Künneth theorem**, since the prototype was proved by D. Künneth (*Math. Ann.*, 90 (1923); *ibid.*, 91 (1924)). The present form was given by H. Cartan and S. Eilenberg [8].

The **tensor product** $C \otimes D$ of cochain complexes C and D is defined analogously to that of chain complexes. Given topological spaces X, Y and a ring Λ , a cochain mapping $\mu: \text{Hom}(S(X), \Lambda) \otimes \text{Hom}(S(Y), \Lambda) \rightarrow \text{Hom}(S(X) \otimes S(Y), \Lambda)$ is defined by $(\mu(u \otimes v))(c \otimes d) = u(c)v(d)$, where $u \in \text{Hom}(S_p(X), \Lambda)$, $v \in \text{Hom}(S_q(Y), \Lambda)$, $c \in S_p(X)$, $d \in S_q(Y)$, and $u(c)v(d)$ is understood to be 0 if $(p, q) \neq (s, t)$. We then have the composite $H^p(X; \Lambda) \otimes H^q(Y; \Lambda) \xrightarrow{\mu^*} H^{p+q}(\text{Hom}(S(X) \otimes S(Y), \Lambda)) \xrightarrow{\rho^*} H^{p+q}(X \times Y; \Lambda)$, where ρ is the Alexander-Whitney mapping. For $\xi \in H^p(X; \Lambda)$ and $\eta \in H^q(Y; \Lambda)$, the **cross product** $\xi \times \eta \in H^{p+q}(X \times Y; \Lambda)$ is defined to be $\rho^* \mu_*(\xi \otimes \eta)$. The cohomology cross product satisfies the properties analogous to the homology cross product.

The cup product and the cohomology cross product are given in terms of each other: $\xi \smile \eta = d^*(\xi \times \eta)$, $\xi \times \eta = \pi_1^*(\xi) \smile \pi_2^*(\eta)$, where $d: X \rightarrow X \times X$ is given by $d(x) = (x, x)$, and $\pi_1: X \times Y \rightarrow X$, $\pi_2: X \times Y \rightarrow Y$ are projections.

We have the following **Künneth theorem** for cohomology: If Λ is a principal ideal domain and each $H_q(X; \Lambda)$ is finitely generated over Λ , then there is an exact sequence

$$0 \rightarrow \sum_{p+q=n} H^p(X; \Lambda) \otimes_\Lambda H^q(Y; \Lambda) \xrightarrow{\times} H^n(X \times Y; \Lambda) \rightarrow \sum_{p+q=n+1} \text{Tor}_\Lambda(H^p(X; \Lambda), H^q(Y; \Lambda)) \rightarrow 0,$$

and this sequence is split (\rightarrow 200 Homological Algebra). For $\xi \in H^p(X; \Lambda)$, $\eta \in H^q(Y; \Lambda)$, $\xi' \in H^s(X; \Lambda)$, $\eta' \in H^t(Y; \Lambda)$, the formula $(\xi \times \eta) \smile (\xi' \times \eta') = (-1)^{qs}(\xi \smile \xi') \times (\eta \smile \eta')$ holds. Therefore, if k is a field and $\dim_k H_q(X; k) < \infty$ for each q , then the cohomology ring $H^*(X \times Y; k)$ is determined by the cohomology rings $H^*(X; k)$ and $H^*(Y; k)$.

†Fiber bundles can be considered as generalized product spaces. Let E be the total space of a fiber bundle with base B and fiber F . The following **Leray-Hirsch theorem** (*J. Math. Pures Appl.*, 29 (1950)) asserts that, under certain conditions, the cohomology of E is additively isomorphic to that of $B \times F$: Let Λ be a principal ideal domain, and assume that $H_*(F; \Lambda)$ is free and finitely generated over Λ . Furthermore, assume that there is a homomorphism $\theta: H^*(F; \Lambda) \rightarrow H^*(E; \Lambda)$ such that the composite $H^*(F; \Lambda) \xrightarrow{\theta} H^*(E; \Lambda) \xrightarrow{i_b^*} H^*(p^{-1}(b); \Lambda)$ is an isomorphism for each $b \in B$, where $p: E \rightarrow B$ is the projection and $i_b: p^{-1}(b) \subset E$. Then an isomorphism $\Phi: H^*(B; \Lambda) \otimes_\Lambda H^*(F; \Lambda) \cong H^*(E; \Lambda)$ is given by $\Phi(\xi \otimes \eta) = p^* \xi \smile \theta(\eta)$, where $\xi \in H^*(B; \Lambda)$, $\eta \in H^*(F; \Lambda)$.

A general connection between (co)homology of E and $B \times F$ is given by means of spectral sequences (\rightarrow 148 Fiber Spaces).

K. Cap and Slant Products

There are other products closely related to the cup product or the cross product that involve cohomology and homology together.

Given a topological space X and a ring Λ , the **cap product** $v \frown c \in S_p(X) \otimes \Lambda$ of a **cochain** $v \in \text{Hom}(S_q(X), \Lambda)$ and a **chain** $c = \sum_i \sigma_i \otimes \lambda_i \in S_{p+q}(X) \otimes \Lambda$ is defined by $v \frown c = \sum_i \sigma_i \circ \varepsilon \otimes v(\sigma_i \circ \varepsilon') \lambda_i$, where σ_i are singular $(p+q)$ -simplexes in X , $\lambda_i \in \Lambda$, and $\varepsilon: \Delta^p \rightarrow \Delta^{p+q}$, $\varepsilon': \Delta^q \rightarrow \Delta^{p+q}$ are the mappings used in the definition of cup product. For any $u \in \text{Hom}(S_p(X), \Lambda)$, the formula $\langle u \smile v, c \rangle = \langle u, v \frown c \rangle$ holds. The cap product satisfies $\partial(v \frown c) = (-1)^p \partial v \frown c + v \frown \partial c$, and hence it induces the **cap product** $\xi \frown a \in H_p(X; \Lambda)$ of a **cohomology class** $\xi \in H^q(X; \Lambda)$ and a **homology class** $a \in H_{p+q}(X; \Lambda)$. If Λ is a commutative ring with 1, then the cap product operation is bilinear and satisfies the following properties: $(\xi \smile \xi') \frown a = \xi \frown (\xi' \frown a)$, $f_*(f^* \eta \frown a) = \eta \frown f_*(a)$, $1 \frown a = a$, $(\xi \times \eta) \frown (a \times b) = (-1)^{p(q-t)} (\xi \frown a) \times (\eta \frown b)$, where $\xi \in H^p(X; \Lambda)$, $\xi' \in H^p(X; \Lambda)$, $\eta \in H^q(Y; \Lambda)$, $a \in H_s(X; \Lambda)$, $b \in H_t(Y; \Lambda)$, and $f: X \rightarrow Y$ is a continuous mapping.

Given topological spaces X , Y and a ring Λ , the **slant product** $w/d \in \text{Hom}(S_p(X), \Lambda)$ of a **cochain** $w \in \text{Hom}(S(X) \otimes S(Y), \Lambda)$ and a **chain** $d = \sum_i \tau_i \otimes \lambda_i \in S_q(Y) \otimes \Lambda$ is defined

by $(w/d)(\sigma) = \sum_i (w(\sigma \otimes \tau_i)) \lambda_i$, where σ is a singular p -simplex in X , τ_i are singular q -simplexes in Y , and $\lambda_i \in \Lambda$. The slant operation satisfies $\delta(w/d) = (\delta w)/d - (-1)^p w/\partial d$. Therefore, under the identification $H^*(X \times Y; \Lambda) = H^*(\text{Hom}(S(X) \otimes S(Y), \Lambda))$, it induces the **slant product** $\zeta/b \in H^p(X; \Lambda)$ of a **cohomology class** $\zeta \in H^{p+q}(X \times Y; \Lambda)$ and a **homology class** $b \in H_q(Y; \Lambda)$. For any $a \in H_p(X; \Lambda)$, it holds that $\langle \zeta/b, a \rangle = \langle \zeta, a \times b \rangle$.

Let G , G' and G'' be Abelian groups. Given a homomorphism $G' \otimes G'' \rightarrow G$, we write $g'g''$ for the image of $g' \otimes g'' \in G' \otimes G''$ in G . Then the cap product $\frown: H^q(X; G') \otimes H_{p+q}(X; G'') \rightarrow H_p(X; G)$ can be defined in the same way as before. Similar definitions are valid for the cup product, the cross products for homology and cohomology, and the slant product.

L. Relative Homology

If C' is a subcomplex of a chain complex C , then we have a chain complex $C/C' = \{C_q/C'_q, \bar{\partial}_q\}$, where C_q/C'_q denotes the quotient group and $\bar{\partial}_q$ is induced from ∂_q by passing to the quotient. C/C' is called the **quotient complex** of C by C' .

A **topological pair** (X, A) is composed of a topological space X and its subset A . Given a topological pair (X, A) and an Abelian group G , we have the chain complex $(S(X)/S(A)) \otimes G$ and the cochain complex $\text{Hom}(S(X)/S(A), G)$. The homology group of $(S(X)/S(A)) \otimes G$ is denoted by $H_*(X, A; G)$ and is called the **relative singular homology group of X modulo A with coefficients in G** or the **singular homology group of (X, A) with coefficients in G** . The homology group $H_*(X; G) = H_*(X, \emptyset; G)$ is sometimes called the **absolute homology group**. Similar definitions are made for the cohomology group $H^*(X, A; G)$ of the cochain complex $\text{Hom}(S(X)/S(A), G)$.

A **simplicial pair** (K, L) is composed of a simplicial complex K and its subcomplex L , and a **CW pair** (X, A) is composed of a CW complex X and its subcomplex A . The relative homology group $H_*(K, L; G) = H_*((C(K)/C(L)) \otimes G)$ of a simplicial pair (K, L) is isomorphic to $H_*(|K|, |L|; G)$. For a CW pair (X, A) , there is an isomorphism $H_*(X, A; G) \cong \tilde{H}_*(X/A; G)$. Similar statements hold for the cohomology groups.

A continuous mapping $f: (X, A) \rightarrow (Y, B)$ of topological pairs is a continuous mapping $f: X \rightarrow Y$ such that $f(A) \subset B$. If $f: (X, A) \rightarrow (Y, B)$ is a continuous mapping, then $f_\#: S(X) \rightarrow S(Y)$ sends $S(A)$ to $S(B)$, and hence f induces homomorphisms $f_*: H_*(X, A; G) \rightarrow H_*(Y, B; G)$ and

$f^*: H^*(Y, B; G) \rightarrow H^*(X, A; G)$. Since the properties are analogous, we state them below only for the case of relative homology.

The following six properties are fundamental. (i) $1_* = 1$. (ii) $(g \circ f)_* = g_* \circ f_*$. (iii) Homotopy property: If $f, f': (X, A) \rightarrow (Y, B)$ are \dagger homotopic, then $f_* = f'_*$. (iv) Exactness property: There exists a **homology exact sequence of** $(X, A): \dots \xrightarrow{\partial_*} H_q(A; G) \xrightarrow{i_*} H_q(X; G) \xrightarrow{j_*} H_q(X, A; G) \xrightarrow{\partial_*} H_{q-1}(A; G) \xrightarrow{i_*} \dots$, where $i: A \subset X, j: (X, \emptyset) \subset (X, A)$, and ∂_* sends the homology class of a cycle of $(S(X)/S(A)) \otimes G$ represented by a chain $c \in S(X) \otimes G$ to the homology class of ∂c which is a cycle of $S(A) \otimes G$. ∂_* is called the **boundary homomorphism** or the **connecting homomorphism**. (v) Naturality of ∂_* : For any continuous mapping $f: (X, A) \rightarrow (Y, B)$, it holds that $\partial_* \circ f_* = (f|_A)_* \circ \partial_*$. (vi) Excision property: If U is a subset of X such that the closure \bar{U} is in $\text{Int } A$, then the **excision isomorphism** $H_*(X - U, A - U; G) \cong H_*(X, A; G)$ is induced by inclusion.

The exactness property extends to the **homology exact sequence of a triple** (X, A, B) : $\dots \xrightarrow{\partial_*} H_q(A, B; G) \xrightarrow{i_*} H_q(X, B; G) \xrightarrow{j_*} H_q(X, A; G) \xrightarrow{\partial_*} H_{q-1}(A, B; G) \xrightarrow{i_*} \dots$. A couple $\{X_1, X_2\}$ of subsets of X is said to be **excisive** if $H_*(X_1, X_1 \cap X_2) \cong H_*(X_1 \cup X_2, X_2)$ is induced by inclusion. If $\{X_1, X_2\}$ is excisive, so is $\{X_2, X_1\}$. For example, if $X = \text{Int } X_1 \cup \text{Int } X_2$ or if X_1 and X_2 are subcomplexes of a CW complex, then $\{X_1, X_2\}$ is excisive. If $\{X_1, X_2\}$ and $\{A_1, A_2\}$ are excisive couples such that $A_1 \subset X_1$ and $A_2 \subset X_2$, then we have the **relative Mayer-Vietoris exact sequence**: $\dots \rightarrow H_q(X_1 \cap X_2, A_1 \cap A_2; G) \rightarrow H_q(X_1, A_1; G) + H_q(X_2, A_2; G) \rightarrow H_q(X_1 \cup X_2, A_1 \cup A_2; G) \rightarrow H_{q-1}(X_1 \cap X_2, A_1 \cap A_2; G) \rightarrow \dots$. For the case of relative cohomology, we use terms such as **cohomology exact sequence** and **coboundary homomorphism**, correspondingly.

The universal coefficient theorems are valid for the relative (co)homology groups. Given a homomorphism $G' \otimes G'' \rightarrow G$, if $\{A, B\}$ is excisive in X , then the cup product $\smile: H^p(X, A; G') \otimes H^q(X, B; G'') \rightarrow H^{p+q}(X, A \cup B; G)$ and the cap products $\frown: H^q(X, B; G') \otimes H_{p+q}(X, A \cup B; G'') \rightarrow H_p(X, A; G)$ can be defined. The product $(X, A) \times (Y, B)$ is defined to be the pair $(X \times Y, A \times Y \cup X \times B)$. Given a homomorphism $G' \otimes G'' \rightarrow G$, the homology cross products $\times: H_p(X, A; G') \otimes H_q(Y, B; G'') \rightarrow H_{p+q}((X, A) \times (Y, B); G)$ and the slant products $/: H^{p+q}((X, A) \times (Y, B); G') \otimes H_q(Y, B; G'') \rightarrow H^p(X, A; G)$ can be defined. If $\{A \times Y, X \times B\}$ is excisive, then the cohomology cross products $\times: H^p(X, A; G') \otimes H^q(Y, B; G'') \rightarrow H^{p+q}((X, A) \times (Y, B); G)$ can also be defined. If $\{A \times Y, X \times B\}$ is excisive, the Künneth theorems are valid for the relative (co)homology groups [9, 10, 11].

M. Čech Homology Theory

Another homology theory is commonly used along with the singular theory. The theory was originated by E. Čech (*Fund. Math.*, 19 (1932)) and was modified by C. H. Dowker (*Ann. Math.*, (2) 51 (1950)).

Given a topological space X and an Abelian group G , the **Čech homology group** $\check{H}_*(X; G)$ and the **Čech cohomology group** $\check{H}^*(X; G)$ are defined as follows. We take the family of all \dagger open coverings of X directed by \dagger refinement, and we consider the \dagger nerve $K(\mathcal{U})$ of each open covering \mathcal{U} , that is, the simplicial complex whose simplexes are finite nonempty subsets of \mathcal{U} with nonempty intersection. If \mathcal{U}' is a refinement of \mathcal{U} , then a simplicial mapping $\pi(\mathcal{U}, \mathcal{U}'): K(\mathcal{U}') \rightarrow K(\mathcal{U})$ is obtained by assigning to each $U' \in \mathcal{U}'$ an element $U \in \mathcal{U}$ such that $U' \subset U$. The induced homomorphisms $\pi(\mathcal{U}, \mathcal{U}')_*: H_*(K(\mathcal{U}'); G) \rightarrow H_*(K(\mathcal{U}); G)$ and $\pi(\mathcal{U}, \mathcal{U}')^*: H^*(K(\mathcal{U}); G) \rightarrow H^*(K(\mathcal{U}'); G)$ are independent of the choice of $\pi(\mathcal{U}, \mathcal{U}')$, and we have the \dagger inverse system $\{H_*(K(\mathcal{U}); G), \pi(\mathcal{U}, \mathcal{U}')_*\}$ and the \dagger direct system $\{H^*(K(\mathcal{U}); G), \pi(\mathcal{U}, \mathcal{U}')^*\}$. We now define $\check{H}_*(X; G) = \varprojlim H_*(K(\mathcal{U}); G)$ and $\check{H}^*(X; G) = \varinjlim H^*(K(\mathcal{U}); G)$.

A continuous mapping $f: X \rightarrow Y$ induces homomorphisms $f_*: \check{H}_*(X; G) \rightarrow \check{H}_*(Y; G)$ and $f^*: \check{H}^*(Y; G) \rightarrow \check{H}^*(X; G)$ as follows. If \mathfrak{B} is an open covering of Y then a simplicial mapping $f_{\mathfrak{B}}: K(f^{-1}(\mathfrak{B})) \rightarrow K(\mathfrak{B})$ is defined by $f_{\mathfrak{B}}(f^{-1}(V)) = V$ ($V \in \mathfrak{B}$). The induced homomorphisms $f_{\mathfrak{B}*}: H_*(K(f^{-1}(\mathfrak{B})); G) \rightarrow H_*(K(\mathfrak{B}); G)$ for all open coverings \mathfrak{B} of Y gives rise to $f_*: \check{H}_*(X; G) \rightarrow \check{H}_*(Y; G)$. Similarly f induces f^* .

Another approach to Čech cohomology theory is called the Alexander-Kolmogorov construction (*Proc. Nat. Acad. Sci.*, 21 (1935) and *C. R. Acad. Sci. Paris*, 202 (1936)). The approach was improved by E. H. Spanier (*Ann. Math.*, (2) 49 (1948)), and the theory is now called the **Alexander- (or Alexander-Kolmogorov-Spanier) cohomology theory**.

The **Alexander cohomology group** $\bar{H}^*(X; G)$ is defined as follows. Let $\Phi^q(X; G)$ be the Abelian group of all functions from the $(q+1)$ -fold product space X^{q+1} to G with addition defined pointwise. An element $\varphi \in \Phi^q(X; G)$ is said to be locally zero if there is an open covering \mathcal{U} of X such that $\varphi(x_0, \dots, x_q)$ vanishes if x_0, \dots, x_q are contained simultaneously in some $U \in \mathcal{U}$. The subgroup of $\Phi^q(X; G)$ consisting of locally zero functions is denoted by $\Phi_0^q(X; G)$. We define a homomorphism $\delta^q: \Phi^q(X; G) \rightarrow \Phi^{q+1}(X; G)$ by $(\delta^q \varphi)(x_0, x_1, \dots, x_{q+1}) = \sum_{i=0}^{q+1} (-1)^i \varphi(x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_{q+1})$. Then $\Phi(X; G) = \{\Phi^q(X; G), \delta^q\}$ is a cochain complex, and $\Phi_0(X; G) = \{\Phi_0^q(X; G), \delta^q\}$ is its subcomplex. We now define $\bar{H}^*(X; G)$ to be

the cohomology group of the quotient complex $\Phi(X; G) = \Phi(X; G)/\Phi_0(X; G)$.

If $f: X \rightarrow Y$ is a continuous mapping, then a cochain map $f^\#: \Phi(Y; G) \rightarrow \Phi(X; G)$ is defined by $(f^\# \varphi)(x_0, x_1, \dots, x_q) = \varphi(f(x_0), f(x_1), \dots, f(x_q))$, and it induces the homomorphism $f^*: \bar{H}^*(Y; G) \rightarrow \bar{H}^*(X; G)$. There is a natural isomorphism $\check{H}^*(X; G) \cong \bar{H}^*(X; G)$.

The (co)homology group of a simplicial complex K is isomorphic to the Čech (co)homology group of $|K|$. If X is a manifold or a CW complex, then its singular (co)homology group and its Čech (co)homology group are isomorphic. However, even for compact metric spaces X , the singular (co)homology group of X is not necessarily isomorphic with the Čech (co)homology group of X .

If $\{X_\lambda\}$ is an \dagger inverse system of compact Hausdorff spaces and $X = \lim X_\lambda$, then there are isomorphisms $\lim \check{H}_*(X_\lambda; G) \cong \check{H}_*(X; G)$ and $\lim \check{H}^*(X_\lambda; G) \cong \check{H}^*(X; G)$. This is called the **continuity property for Čech theory**. If A is any closed subset of a manifold M , then there is an isomorphism $\lim H^*(W; G) \cong \check{H}^*(A; G)$, where W varies over neighborhoods of A in M directed downward by inclusion. If the \dagger covering dimension of X is n , $\check{H}^q(X; G) = 0$ for $q > n$.

The cup product in the Čech cohomology is introduced simply by passing to the limit with cup products in simplicial complexes, and the cup product in the Alexander cohomology is induced from the operation $\smile: \Phi^p(X; G') \otimes \Phi^q(X; G'') \rightarrow \Phi^{p+q}(X; G)$ defined by $(\varphi \smile \psi)(x_0, x_1, \dots, x_{p+q}) = \varphi(x_0, x_1, \dots, x_p)\psi(x_p, x_{p+1}, \dots, x_{p+q})$.

The **relative Čech homology group** $\check{H}_*(X, A; G)$ and the **relative Čech cohomology group** $\check{H}^*(X, A; G)$ are defined as follows: An open covering of (X, A) is a pair $(\mathcal{U}, \mathcal{V})$ of an open covering \mathcal{U} of X and an open covering \mathcal{V} of A such that $\mathcal{V} \subset \mathcal{U}$. To such a pair $(\mathcal{U}, \mathcal{V})$ we assign a simplicial pair $(K(\mathcal{U}), K'(\mathcal{V}))$, where $K'(\mathcal{V})$ is the nerve of $\mathcal{V} \cap A = \{N \cap A \mid N \in \mathcal{V}\}$. Considering the family of all open coverings of (X, A) , we define now $\check{H}_*(X, A; G) = \lim H_*(K(\mathcal{U}), K'(\mathcal{V}); G)$ and $\check{H}^*(X, A; G) = \varprojlim H^*(K(\mathcal{U}), K'(\mathcal{V}); G)$.

The **relative Alexander cohomology group** $\bar{H}^*(X, A; G)$ is defined to be the cohomology group of the cochain complex which is the kernel of the cochain mapping $i^\#: \Phi(X; G) \rightarrow \Phi(A; G)$ induced by inclusion. There is a natural isomorphism $\check{H}^*(X, A; G) \cong \bar{H}^*(X, A; G)$.

The relative Čech (co)homology groups satisfy the properties analogous to the relative singular (co)homology groups except the exactness property for homology (\rightarrow Section Q). In certain cases, the excision property is strengthened for Čech (co)homology. For example, we have the following theorem: Assume that X and Y are compact Hausdorff

spaces, A and B are closed subsets of X and Y , respectively, and $f: (X, A) \rightarrow (Y, B)$ is a continuous mapping which maps $X - A$ onto $Y - B$ homeomorphically. Then $f_*: \check{H}_*(X, A; G) \cong \check{H}_*(Y, B; G)$ and $f^*: \check{H}^*(Y, B; G) \cong \check{H}^*(X, A; G)$ hold.

N. Fundamental Classes of Manifolds

For a topological space X and a point x of X , the **local homology groups** $H_*(X, X - x)$ represent a topological property of X around x .

The notion of \dagger orientation for differentiable manifolds and triangulable manifolds generalizes to \dagger topological manifolds by using local homology groups as follows. Let M be an n -dimensional (topological) manifold with \dagger boundary ∂M . If x is a point of the interior $M_0 = M - \partial M$, then $H_q(M, M - x) \cong H_q(\mathbf{R}^n, \mathbf{R}^n - 0)$ is \mathbf{Z} for $q = n$ and is 0 for $q \neq n$. We define a **local orientation** o_x for M at $x \in M_0$ to be a choice of one of the two possible generators for $H_n(M, M - x)$, and we then define an **orientation** for M to be a function which assigns to each $x \in M_0$ a local orientation o_x which varies continuously with x in the following sense: For each x there should exist a compact neighborhood $N \subset M_0$ and an element $o_N \in H_n(M, M - N)$ such that $i_{y*}(o_N) = o_y$ for each $y \in N$, where $i_y: (M, M - N) \subset (M, M - y)$. If there is an orientation for M , then M is said to be **orientable**, and the pair of M and an orientation is called an **oriented manifold**. If M is a nonorientable manifold without boundary, the set of local orientations for M forms an orientable manifold doubly covering M , called the **orientation manifold** of M .

If M is an oriented n -dimensional manifold, then for any compact subset K of M there is a unique element $o_K \in H_n(M, (M - K) \cup \partial M)$ such that $i_{x*}(o_K) = o_x$ for each $x \in K \cap M_0$, where $i_x: (M, (M - K) \cup \partial M) \subset (M, M - x)$. The element o_K is called the **fundamental homology class around K** . In particular, if M is itself compact, $o_M \in H_n(M, \partial M)$ is usually denoted by $[M]$ and is called the **fundamental homology class** of M . A connected compact n -dimensional manifold M is orientable if and only if $H_n(M, \partial M) \neq 0$, and in this case $H_n(M, \partial M)$ is a free cyclic group generated by a fundamental class $[M]$. If M is an orientable compact n -dimensional manifold, then ∂M is an orientable compact $(n - 1)$ -dimensional manifold without boundary, and the boundary homomorphism $\partial_*: H_n(M, \partial M) \rightarrow H_{n-1}(\partial M)$ sends a fundamental class $[M]$ to a fundamental class $[\partial M]$.

An n -dimensional manifold M is orientable if and only if there exists an element $U \in H^n(M \times M, M \times M - dM)$ such that, for each $x \in M_0$,

$j_x^*(U)$ is a generator of $H^n(M, M-x)$, where dM is the diagonal in $M \times M$, and $j_x: (M, M-x) \rightarrow (M \times M, M \times M - dM)$ is given by $j_x(y) = (x, y)$ ($y \in M$). In fact, U corresponds to an orientation which assigns to each $x \in M_0$ a local orientation o_x such that $\langle j_x^*(U), o_x \rangle = 1$. The element U is called the **orientation cohomology class** of M . If M is a compact manifold without boundary, it holds that $\langle d^*(U), [M] \rangle = \chi(M)$, where $d^*: H^n(M \times M, M \times M - dM) \rightarrow H^n(M)$ is induced by the diagonal mapping. The element $d^*(U) \in H^n(M)$ is called the **Euler class** of M .

If we work with the (co)homology groups with coefficients in \mathbb{Z}_2 , the fundamental classes are defined for an arbitrary manifold without making any assumption of orientability. If M is connected and compact, then $H_n(M, \partial M; \mathbb{Z}_2) \cong \mathbb{Z}_2$ is generated by $[M]$.

O. Duality in Manifolds

Let M be a compact n -dimensional manifold, and let M_1, M_2 be compact $(n-1)$ -dimensional manifolds such that $M_1 \cup M_2 = \partial M$ and $M_1 \cap M_2 = \partial M_1 = \partial M_2$. Assume either that M is oriented or that $G = \mathbb{Z}_2$. Then for each q , an isomorphism $D: H^q(M, M_1; G) \cong H_{n-q}(M, M_2; G)$ is defined by $D(\xi) = \xi \frown [M]$. In particular, there are the isomorphisms $D: H^q(M, \partial M; G) \cong H_{n-q}(M; G)$ and $H^q(M; G) \cong H_{n-q}(M, \partial M; G)$, where the cap product is taken with respect to the homomorphism $G \otimes \mathbb{Z} \rightarrow G$ defined by multiplication. This theorem is called the **Poincaré-Lefschetz duality theorem**, and the special case for $\partial M = \emptyset$ is often referred to as **Poincaré duality**.

Poincaré duality implies the following consequences for a compact n -dimensional manifold M without boundary. If M is orientable, then the q th Betti number is equal to the $(n-q)$ th Betti number, and the q th torsion coefficients are equal to the $(n-q-1)$ th torsion coefficients. If n is odd, then $\chi(M) = 0$, and if M is orientable and $n \equiv 2 \pmod{4}$, then $\chi(M)$ is even.

Poincaré duality generalizes to the following duality theorem. Let M be an n -dimensional manifold without boundary, and let K be a compact subset of M . Assume either that M is oriented or that $G = \mathbb{Z}_2$. Then there is an isomorphism $D: \tilde{H}^q(K; G) \cong H_{n-q}(M, M-K; G)$ for any q , which is given as follows: For each open neighborhood W of K , define $D_W: H^q(W; G) \rightarrow H_{n-q}(M, M-K; G)$ by $D_W(\xi) = k_*(\xi \frown k_*^{-1}(o_K))$, where k_* is the excision isomorphism induced by $k: (W, W-K) \hookrightarrow (M, M-K)$. Now D is defined to be the limit of D_W , where W varies over open neighborhoods of K . The inverse of D up to sign is given in terms of the slant

product with the fundamental cohomology class U of M as follows: For each open neighborhood W of K , we define a homomorphism $\gamma_W: H_{n-q}(M, M-W; G) \rightarrow H^q(W; G)$ by $\gamma_W(a) = j^*(U)/a$, where $j^*: H^n(M \times M, M \times M - dM) \rightarrow H^n(W \times (M, M-W))$ is induced by inclusion. Then, passing to the limit, these γ_W define the desired one.

If we take an n -sphere S^n as M in the above duality theorem and use the homology exact sequence of $(S^n, S^n - K)$, then we have the following **Alexander duality theorem** (*Trans. Amer. Math. Soc.*, 23 (1922)): If K is a closed subset of S^n , then the q th reduced Čech cohomology group of K is isomorphic to the $(n-q-1)$ th reduced singular homology group of $S^n - K$ for any coefficient group G and any q . In particular, if K is a neighborhood retract, $\tilde{H}^q(K; G) \cong \tilde{H}_{n-q-1}(S^n - K; G)$ holds. This shows that $H_*(S^n - K)$ depends only on K and not on the way K is embedded in S^n . The Alexander duality theorem for $n=2$ and $K=S^1$ gives the classical \dagger Jordan curve theorem.

In view of the duality theorems, certain classical definitions in the homology of manifolds can be given in terms of cohomology. For example, if $f: M \rightarrow M'$ is a continuous mapping of oriented closed manifolds, then the **Umkehr homomorphism** or **Gysin homomorphism** $f^!: H_q(M'; G) \rightarrow H_{q+d}(M; G)$ ($d = \dim M - \dim M'$) (W. Gysin, *Comment. Math. Helv.*, 14 (1941)) can be defined by $D \circ f^! = f^* \circ D$. In cohomology we have $f!: H^q(M; G) \rightarrow H^{q-d}(M'; G)$. Similarly, if M is an oriented n -dimensional closed manifold and $a \in H_p(M)$, $b \in H_q(M)$, then the **intersection product** $a \cdot b \in H_{p+q-n}(M)$ of Lefschetz can be defined by $a \cdot b = D^{-1}a \frown b = D(D^{-1}a \smile D^{-1}b)$. If $p+q=n$, the number $a \cdot b \in H_0(M) \cong \mathbb{Z}$ is called the **intersection number** of a and b . The classical definitions are still meaningful today, since they are closer to geometric intuition and therefore possess considerable heuristic value. For example, the following fact serves to compute cup products in manifolds. If M is an oriented closed differentiable manifold and $a, b \in H_*(M)$ are represented by closed submanifolds N_1, N_2 which intersect \dagger transversally, then $\pm a \cdot b$ is represented by $N_1 \cap N_2$ [11]. See [12] for a rigorous discussion of classical intersection theory.

P. Cohomology with Compact Supports

Let X be a topological space. A subset V of X is said to be **cobounded** if $\overline{X-V}$ is compact. A singular q -cochain $u \in \text{Hom}(S_q(X), G)$ is said to have **compact support** if there exists a cobounded set V such that $u(\sigma) = 0$ for every singular q -simplex σ in V . The singular cochains with compact support form a subcom-

plex of the cochain complex $\text{Hom}(S(X), G)$. The cohomology group of this subcomplex is denoted by $H_c^*(X; G)$ and is called the **singular cohomology group of X with compact supports**. There is an isomorphism $H_c^*(X; G) \cong \varinjlim H^*(X, V; G)$, where V varies over cobounded subsets of X .

Let K be a simplicial complex. A q -cochain $u \in \text{Hom}(C_q(K), G)$ is called a **finite cochain** of K if $u(\sigma) = 0$ except for a finite number of oriented q -simplexes σ of K . If K is a \dagger locally finite simplicial complex, then finite cochains of K form a subcomplex of the cochain complex $\text{Hom}(C(K), G)$ whose cohomology group is isomorphic to $H_c^*(|K|; G)$.

Let X be a \dagger locally compact Hausdorff space, and let $X \cup \{\infty\}$ denote the \dagger one-point compactification of X . Then the **Čech cohomology group of X with compact supports**, denoted by $\check{H}_c^*(X; G)$ is defined to be the reduced Čech cohomology group of $X \cup \{\infty\}$ with coefficients in G . There is an isomorphism $\check{H}_c^*(X; G) \cong \varinjlim \check{H}^*(X, V; G)$, where V varies over cobounded subsets of X . If X is a manifold or a CW complex, then $H_c^*(X; G) \cong \check{H}_c^*(X; G)$. If X is a compact Hausdorff space and A is closed in X , then $\check{H}_c^*(X - A; G) \cong H^*(X, A; G)$. The Alexander-Kolmogorov construction gives a direct approach to $\check{H}_c^*(X; G)$ [10, 13].

A \dagger proper continuous mapping $f: X \rightarrow Y$ of locally compact Hausdorff spaces induces homomorphisms $f^*: H_c^*(Y; G) \rightarrow H_c^*(X; G)$ and $f^*: \check{H}_c^*(Y; G) \rightarrow \check{H}_c^*(X; G)$, and if $f, f': X \rightarrow Y$ are properly homotopic, then they induce the same homomorphisms.

The cohomology with compact supports is useful in order to extend results in the cohomology of compact spaces to noncompact spaces. For example, the conclusion of the duality theorem on a compact set $K \subset M$ in Section O generalizes to the case of a closed set $K \subset M$ as follows: There is an isomorphism $\check{H}_c^q(K; G) \cong H_{n-q}(M, M - K; G)$ for any q . This implies the following generalization of Poincaré duality: $H_c^q(M; G) \cong H_{n-q}(M; G)$ holds for an orientable n -dimensional manifold M without boundary.

There are homology theories associated with the cohomology theories with compact supports [13].

Q. Eilenberg-Steenrod Axioms

Let H^* be a collection of the following three functions: (1) A function assigning to each topological pair (X, A) and each integer q an Abelian group $H^q(X, A)$. (2) A function assigning to each continuous mapping $f: (X, A) \rightarrow (Y, B)$ and each integer q a homomorphism

$f^*: H^q(Y, B) \rightarrow H^q(X, A)$. (3) A function assigning to each topological pair (X, A) and each integer q a homomorphism $\delta^*: H^q(A) \rightarrow H^{q+1}(X, A)$. Then H^* is called a **cohomology theory on the category of topological pairs** if the following seven axioms are satisfied [7]. (i) $1_* = 1$, where 1 is identity. (ii) $(g \circ f)^* = f^* \circ g^*: H^q(Z, C) \rightarrow H^q(X, A)$ for continuous mappings $f: (X, A) \rightarrow (Y, B)$ and $g: (Y, B) \rightarrow (Z, C)$. (iii) Homotopy axiom: If $f, f': (X, A) \rightarrow (Y, B)$ are homotopic, then $f_* = f'_*: H^q(Y, B) \rightarrow H^q(X, A)$. (iv) Exactness axiom: The sequence $\dots \xrightarrow{\delta^*} H^q(X, A) \xrightarrow{i^*} H^q(X) \xrightarrow{i^*} H^q(A) \xrightarrow{\delta^*} H^{q+1}(X, A) \xrightarrow{j^*} \dots$ is exact, where $i: A \subset X$ and $j: (X, \emptyset) \subset (X, A)$. (v) $f^* \circ \delta^* = \delta^* \circ (f|A)^*: H^q(B) \rightarrow H^{q+1}(X, A)$ for a continuous mapping $f: (X, A) \rightarrow (Y, B)$. (vi) Excision axiom: If U is an open set of X such that $\bar{U} \subset \text{Int } A$, then $i^*: H^q(X, A) \cong H^q(X - U, A - U)$, where i is the inclusion. (vii) Dimension axiom: $H^q(pt) = 0$ if $q \neq 0$. Axioms (i)–(vii) are called the **Eilenberg-Steenrod axioms**, and the group $H^0(pt)$ is called the **coefficient group** of the cohomology theory H^* .

A cohomology theory on the category of pairs of compact Hausdorff spaces is defined similarly. A cohomology theory on the category of CW pairs (or finite CW pairs) is defined similarly except that axiom (vi) is replaced by the following excision axiom: If $\{X_1, X_2\}$ is a couple of subcomplexes of a CW complex, then $i^*: H^q(X_1 \cup X_2, X_2) \cong H^q(X_1, X_1 \cap X_2)$, where i is the inclusion. Two cohomology theories H^* and H'^* on the same category are **isomorphic** if there is an isomorphism $h_q: H^q(X, A) \cong H'^q(X, A)$ for each (X, A) and each q , and they commute with f^* and δ^* . A homology theory on various categories is defined similarly by dualization.

A singular (co)homology theory with coefficients in G is an example of a (co)homology theory on the category of topological pairs. The Čech cohomology groups with coefficients in G can be made into a cohomology theory on the category of topological pairs. However, the Čech homology groups do not constitute a homology theory on the category of topological pairs; the homology sequence of any pair (X, A) is defined, but it can be proved only that the composite of any two successive homomorphisms is zero. The Čech homology groups with coefficients in a field constitute a homology theory on the category of compact Hausdorff pairs. The Alexander cohomology groups constitute a cohomology theory on the same category of topological pairs, and it is isomorphic to the Čech cohomology theory if their coefficient groups are isomorphic [10].

The Čech (co)homology constitutes a (co)homology theory on the category of CW pairs, and it is isomorphic to the singular (co)homology theory on the same category if

the coefficient groups are isomorphic. (Co-)homology theories on the category of finite CW pairs are determined, up to isomorphisms, by their coefficient groups. This fact is called the **uniqueness theorem of homology theory** on the category of finite CW pairs. Cohomology theories on the category of pairs of compact Hausdorff spaces which satisfy the following continuity axiom are determined, up to isomorphisms, by their coefficient groups: If $\{(X_\lambda, A_\lambda)\}$ is an inverse system of pairs of compact Hausdorff spaces, then $H^q(\varprojlim X_\lambda, \varprojlim A_\lambda) \cong \varprojlim H^q(X_\lambda, A_\lambda)$. The Čech cohomology theory satisfies this axiom.

During recent years, many (co)homology theories have been developed which satisfy the first six Eilenberg-Steenrod axioms but fail to satisfy the dimension axiom. These are called **generalized (co)homology theories**, and include various $\dagger K$ -theories, \dagger bordism theories, and \dagger stable homotopy theories (\rightarrow 202 Homotopy Theory).

R. Homology with Coefficients in Local Systems

N. E. Steenrod (*Ann. Math.*, (2) 44 (1943)) introduced the (co)homology group with coefficients in a local system of Abelian groups, which is useful in \dagger obstruction theory and in the homology theory of \dagger fiber spaces.

A **local system** \mathfrak{G} of Abelian groups on a topological space X is a set of Abelian groups G_x , one for each $x \in X$, together with an isomorphism $l^*: G_{l(0)} \rightarrow G_{l(1)}$ for each \dagger path $l: [0, 1] \rightarrow X$ subject to the following conditions: (1) If two paths l and l' are homotopic with endpoints fixed, then $l^* = l'^*$. (2) If l and m are paths such that $l(1) = m(0)$, then $(l \cdot m)^* = m^* \circ l^*$, where $l \cdot m$ denotes the \dagger product of l and m . An example is provided by the \dagger homotopy groups $\pi_n(X, x)$ for $n \geq 2$. Let M be an n -dimensional topological manifold. Then $x \rightarrow H_n(M, M - x)$ is a local system of infinite cyclic groups. It is called the **orientation sheaf** of M . A local system \mathfrak{G} is said to be trivial if $l^* = l'^*$ for any paths l, l' with the same initial and final points.

Given a local system \mathfrak{G} of Abelian groups on a topological space X , a chain complex $S(X; \mathfrak{G}) = \{S_q(X; \mathfrak{G}), \partial_q\}$ is defined as follows: If $q < 0$, then $S_q(X; \mathfrak{G}) = 0$, and if $q \geq 0$, then $S_q(X; \mathfrak{G})$ is the Abelian group of formal finite sums $\sum g_\sigma \sigma$, where $\sigma: \Delta^q \rightarrow X$ are singular q -simplexes in X and $g_\sigma \in G_{\sigma(1,0,\dots,0)}$; the boundary operator $\partial_q: S_q(X; \mathfrak{G}) \rightarrow S_{q-1}(X; \mathfrak{G})$ is given by $\partial_q(g_\sigma \sigma) = l_\sigma^*(g_\sigma) \sigma \circ \varepsilon_0 + \sum_{i=1}^q (-1)^i g_\sigma \sigma \circ \varepsilon_i$, where $\sigma \circ \varepsilon_i$ is the i th face of σ , and $l_\sigma: [0, 1] \rightarrow X$ is given by $l_\sigma(t) = \sigma(1-t, t, 0, \dots, 0)$. The homology group of the chain complex $S(X; \mathfrak{G})$

is denoted by $H_*(X; \mathfrak{G})$ and is called the **singular homology group of X with coefficients in \mathfrak{G}** .

Similarly, a cochain complex $S^*(X; \mathfrak{G}) = \{S^q(X; \mathfrak{G}), \delta^q\}$ is defined as follows: If $q < 0$, then $S^q(X; \mathfrak{G}) = 0$, and if $q \geq 0$, then $S^q(X; \mathfrak{G})$ is the Abelian group of functions u assigning to every singular q -simplex σ in X an element $u(\sigma) \in G_{\sigma(1,0,\dots,0)}$; the coboundary operator $\delta^q: S^q(X; \mathfrak{G}) \rightarrow S^{q+1}(X; \mathfrak{G})$ for $q \geq 0$ is given by $(\delta^q u)(\sigma) = l_\sigma^{*-1} u(\sigma \circ \varepsilon_0) + \sum_{i=1}^{q+1} (-1)^i u(\sigma \circ \varepsilon_i)$, where σ is a singular $(q+1)$ -simplex in X . The cohomology group of the cochain complex $S^*(X; \mathfrak{G})$ is denoted by $H^*(X; \mathfrak{G})$ and is called the **singular cohomology group of X with coefficients in \mathfrak{G}** .

If \mathfrak{G} is trivial, then the (co)homology group with coefficients in \mathfrak{G} coincides with the (co-)homology group with coefficients in $G \cong G_x$. The various notions and theorems on the ordinary (co)homology can be extended to (co)homology with coefficients in \mathfrak{G} . The Čech (co)homology group with coefficients in \mathfrak{G} is also defined [10]. The cohomology groups with coefficients in \mathfrak{G} are generalized to the cohomology groups with coefficients in a sheaf [10, 14] (\rightarrow 383 Sheaves).

References

- [1] H. Poincaré, *Analysis situs*, J. École Polytech., 1 (1895), 1–121; *Rend. Circ. Mat. Palermo*, 13 (1899), 285–343.
- [2] O. Veblen, *Analysis situs*, *Amer. Math. Soc. Colloq. Publ.*, 1922.
- [3] S. Lefschetz, *Topology*, *Amer. Math. Soc. Colloq. Publ.*, 1930.
- [4] H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Teubner, 1934 (Academic Press, 1980).
- [5] P. S. Aleksandrov and H. Hopf, *Topologie I*, Springer, 1935 (Chelsea, 1965).
- [6] S. Lefschetz, *Algebraic topology*, *Amer. Math. Soc. Colloq. Publ.*, 1942.
- [7] S. Eilenberg and N. E. Steenrod: *Foundations of algebraic topology*, Princeton Univ. Press, 1952.
- [8] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton Univ. Press, 1956.
- [9] P. J. Hilton and S. Wylie, *Homology theory*, Cambridge Univ. Press, 1960.
- [10] E. H. Spanier, *Algebraic topology*, McGraw-Hill, 1966.
- [11] A. Dold, *Lectures on algebraic topology*, Springer, 1972.
- [12] M. Glezerman and L. S. Pontryagin, *Intersections in manifolds*, *Amer. Math. Soc. Transl.*, 50 (1951). (Original in Russian, 1947.)
- [13] W. S. Massey, *Homology and cohomology theory*, Dekker, 1978.

- [14] R. Godement, *Topologie algébrique et théorie des faisceaux*, Hermann, 1958.
- [15] S. Lefschetz, *Introduction to topology*, Princeton Univ. Press, 1949.
- [16] J. G. Hocking and G. Young, *Topology*, Addison Wesley, 1961.
- [17] D. G. Bourgin, *Modern algebraic topology*, Macmillan, 1963.
- [18] H. Schubert, *Topologie*, Teubner, 1964 (Allyn and Bacon, 1968).
- [19] S. T. Hu, *Homology theory*, Holden-Day, 1966.
- [20] M. J. Greenberg, *Lectures on algebraic topology*, Benjamin, 1967.
- [21] G. E. Cooke and R. L. Finney, *Homology of cell complexes*, Princeton Univ. Press, 1967.
- [22] C. R. F. Maunder, *Algebraic topology*, Van Nostrand, 1970.
- [23] A. H. Wallace, *Algebraic topology*, Benjamin, 1970.
- [24] C. Godbillon, *Elements de topologie algébrique*, Hermann, 1971.
- [25] J. Vick, *Homology theory*, Academic Press, 1973.
- [26] J. W. Milnor and J. D. Stasheff, *Characteristic classes*, Princeton Univ. Press, 1974.
- [27] W. S. Massey, *Singular homology theory*, Springer, 1980.

202 (IX.9) Homotopy Theory

A. General Remarks

Given a topological space X , we utilize the concept of homotopy to define the †fundamental group, homotopy groups, and cohomotopy groups of X . These groups, together with (co)homology groups, are useful tools in topology.

Since the research of H. Hopf, W. Hurewicz, and H. Freudenthal in the 1930s, homotopy theory has made rapid progress and now plays an important role in topology.

B. Homotopy

If a family $f_t: X \rightarrow Y$ ($t \in I = \{t | 0 \leq t \leq 1\}$) of †continuous mappings from a †topological space X into a topological space Y is also continuous with respect to t , that is, if the mapping F from the product space $X \times I$ into Y defined by $F(x, t) = f_t(x)$ ($x \in X, t \in I$) is continuous, then $\{f_t\}$ or F is called a **homotopy**. In this case, f_0 and f_1 are said to be **homotopic**. This relation between f_0 and f_1 is indicated by $f_0 \simeq f_1: X \rightarrow Y$, or simply $f_0 \simeq f_1$, and is called

the relation of homotopy. Denote by Y^X the set of all continuous mappings from X into Y . The homotopy relation is an †equivalence relation on Y^X , and the equivalence class $[f]$ of a mapping $f: X \rightarrow Y$ is called the **homotopy class** (or **mapping class**) of f . The set of all homotopy classes of mappings of X into Y is called the **homotopy set** and is denoted by $\pi(X; Y)$ or $[X, Y]$. A function γ of continuous mappings $f \in Y^X$ is called a **homotopy invariant** if $f \simeq g$ implies $\gamma(f) = \gamma(g)$. When X consists of a point $*$ we write $\pi(*; Y) = \pi_0(Y)$. If all continuous mappings in Y^X are homotopic to each other, we write $\pi(X; Y) = 0$; $\pi_0(Y) = 0$ means that Y is †arcwise connected. A mapping f from a compact space into an n -dimensional sphere S^n is called essential if any mapping g homotopic to f satisfies $g(X) = S^n$. A mapping is inessential if and only if it is homotopic to the constant mapping.

These concepts are generalized as follows: Let A_i and B_i ($i = 1, 2, \dots$) be subspaces of X and Y , respectively, and denote by $Y^X(A_1, A_2, \dots; B_1, B_2, \dots)$ the set of continuous mappings $f \in Y^X$ satisfying $f(A_i) \subset B_i$. If a homotopy $\{f_t\}$ is such that $f_t \in Y^X(A_i; B_i)$, then $\{f_t\}$ is called a **restricted homotopy** with respect to A_i, B_i or a homotopy from a system of spaces (X, A_1, A_2, \dots) into a system of spaces (Y, B_1, B_2, \dots) . The notation $f_0 \simeq f_1: (X, A_1, A_2, \dots) \rightarrow (Y, B_1, B_2, \dots)$ and the homotopy set $\pi(X, A_1, A_2, \dots; Y, B_1, B_2, \dots)$ are defined accordingly.

For the composite $g \circ f \in Z^X(A_i; C_i)$ of $f \in Y^X(A_i; B_i)$ and $g \in Z^X(B_i; C_i)$, $f \simeq f'$ and $g \simeq g'$ imply $g \circ f \simeq g' \circ f'$. Thus the **composite** $\beta \circ \alpha = [g \circ f] \in \pi(X, A_i; Z, C_i)$ of $[f] = \alpha \in \pi(X, A_i; Y, B_i)$ and $[g] = \beta \in \pi(Y, B_i; Z, C_i)$ is defined. By putting $g_*[f] = [g \circ f] = f^*[g]$ we induce two mappings,

$$g_*: \pi(X, A_i; Y, B_i) \rightarrow \pi(X, A_i; Z, C_i),$$

$$f^*: \pi(Y, B_i; Z, C_i) \rightarrow \pi(X, A_i; Z, C_i).$$

Then $f \simeq f'$ implies $f^* = f'^*$ and $g \simeq g'$ implies $g_* = g'_*$. Also $(g \circ f)_* = g_* \circ f_*$, $(g \circ f)^* = f^* \circ g^*$, and $h^* \circ g_* = g_* \circ h^*$, where $h \in X^W(D_i; A_i)$.

The **category of pointed topological spaces** is defined to be the †category in which each object X , which is a topological space, has a point fixed as a **base point** and each mapping $X \rightarrow Y$ carries the base point of X to the base point of Y . In this category, we define a homotopy set, denoted by $\pi(X; Y)_0$ or $[X; Y]_0$, as follows: Denoting the base points by $*$, we have $\pi(X, A_i, *; Y, B_i, *) = \pi(X, A_i; Y, B_i)_0$. A continuous mapping f homotopic to the constant mapping $X \rightarrow * \in Y$ is said to be **homotopic to zero** (or **null-homotopic**). This is indicated by $f \simeq 0$, and $\pi(X; Y)_0 = 0$ means that all continuous mappings are homotopic to zero. Let S^0 be a set of two points; then

$\pi(X; S^0)_0 = 0$ means that X is \dagger connected. In contrast to these specific homotopies, the usual homotopy is sometimes called a **free homotopy**.

Suppose that a homotopy $\{f_t\}$ ($f_t: X \rightarrow Y$) is such that the restriction of f_t to a subspace A of X is stationary, that is, $f_t(a) = f_0(a)$ ($a \in A$, $t \in I$). Then f_0, f_1 are said to be **homotopic relative to A** , indicated by $f_0 \simeq f_1$ (rel. A). If a homotopy $\{f_t\}$ ($f_t: X \rightarrow Y$) is such that each f_t is a homeomorphism into Y , then $\{f_t\}$ is called an **isotopy** and f_0 is called **isotopic** to f_1 (\rightarrow 235 Knot Theory).

Research done by L. E. Brouwer, H. Hopf, W. Hurewicz, K. Borsuk, L. S. Pontryagin, and S. Eilenberg has contributed to the theory of homotopy, an important field of topology still in the process of development.

C. Mapping Spaces

We endow the set Y^X of all continuous mappings $f: X \rightarrow Y$, with the \dagger compact-open topology. The topological space Y^X is called a **mapping space**. In particular we denote Y^I ($0, 1; *, *$) ($* \in Y$) by $\Omega(Y) = \Omega(Y, *)$ and call it the **space of closed paths** (or **loop space**) of Y . Two points f, g of Y^X are connected by a \dagger path in Y^X if and only if $f \simeq g: X \rightarrow Y$. Thus $\pi_0(Y^X) = \pi(X; Y)$ and $\pi_0(Y^X(A_i; B_i)) = \pi(X, A_i; Y, B_i)$.

D. Retracts

Let A be a subspace of a topological space X . If there exists an $f \in A^X$ such that the restriction $f|_A$ is the identity mapping of A , then A is called a **retract** of X , and f a **retraction**. If A is a retract of X , any continuous mapping of A into any topological space can be extended to a continuous mapping of X . If A is a retract of some neighborhood $U(A)$, A is called a **neighborhood retract** or **NR** of X . If for any \dagger homeomorphism of a metric space A onto a closed subspace A_0 of any metric space X , A_0 is a retract (neighborhood retract) of X , then A is called an **absolute retract** or **AR** (**absolute neighborhood retract** or **ANR**). For example, an n -dimensional simplex or an n -dimensional Euclidean space is an AR. If a retraction f is homotopic to the identity mapping of X (resp. $U(A)$), we call A a **deformation retract** (**neighborhood deformation retract**) of X . Moreover, if $f \simeq 1_X$ (rel. A), then A is called a **strong deformation retract**. In particular, if a point x_0 is a (strong) deformation retract of X , we say that X is **contractible** to the point x_0 . For example, any \dagger polyhedron P and any compact n -dimensional \dagger topological manifold are ANRs; any polyhedron P_0 contained in P in a strong deformation retract of some neighbor-

hood in P . X is called **locally contractible** if each point x of X has a contractible neighborhood U of x .

E. The Extension Property

Let X, Y be topological spaces, $A \subset X$, $f_0, f_1 \in Y^X$, and $\{g_t: A \rightarrow Y\}$ a homotopy such that $g_t = f_t|_A$ ($t = 0, 1$). We can extend $\{g_t\}$ to a homotopy $\{f_t\}$ of X if and only if the mapping $F: (X \times 0) \cup (A \times I) \cup (X \times 1) \rightarrow Y$ defined by $F(x, i) = f_i(x)$, $F(a, t) = g_t(a)$ can be extended to a continuous mapping sending $X \times I$ into Y . Therefore the problem of whether $f_0 \simeq f_1$ can be reduced to the problem of whether a continuous mapping defined on a subspace can be extended to the whole space. If for any homotopy $\{g_t: A \rightarrow Y\}$ and any continuous mapping $f_0: X \rightarrow Y$ into any topological space Y satisfying $f_0|_A = g_0$ there exists a homotopy $\{f_t: X \rightarrow Y\}$ satisfying $f_t|_A = g_t$, then we say that (X, A) has the **homotopy extension property**. This occurs if and only if $(X \times 0) \cup (A \times I)$ is a retract of $X \times I$. A pair (X, A) of ANRs, where A is closed in X , and a pair (P, P_0) with P a \dagger CW complex and P_0 a subcomplex of P have this property. Given a continuous mapping $h: B \rightarrow A$ of a subspace B of a topological space Y into a topological space A , we identify $b \in B$ with $h(b) \in A$ in the \dagger direct sum $A \cup Y$ and obtain the \dagger identification space denoted by $A \cup_h Y$, which is called an **attaching space** under h . If (Y, B) has the homotopy extension property, then $(Y \times X, B \times X)$ and $(A \cup_h Y, A)$ also have the same property. When A consists of a point $*$, we write $* \cup_h Y = Y/B$ and call the space Y/B a **space smashing** (shrinking or pinching) B to a point. If $Y = B \times I$, $B = B \times 0$, then we call $A \cup_h (B \times I)$ a **mapping cylinder** of h , $(A \cup_h (B \times I))/(B \times 1)$ a **mapping cone** of h , and the mapping cylinder and mapping cone of $h: B \rightarrow *$ the **cone over B** and **suspension** of B , respectively.

F. Homotopy Type

For systems $(X, A_i), (Y, B_i)$ of topological spaces, if there exist $f \in Y^X(A_i; B_i)$, $g \in X^Y(B_i; A_i)$ such that $g \circ f$ and $f \circ g$ are homotopic to the identity mappings of (X, A_i) and (Y, B_i) , respectively, then we say that (X, A_i) and (Y, B_i) have the same **homotopy type** or are **homotopy equivalent**. Such mappings f and g are called **homotopy equivalences**. For a homotopy equivalence f , the induced mappings f_* and f^* are bijective. Therefore, in homotopy theory, systems of spaces having the same homotopy type are considered equivalent. If A is a deformation retract of X , then A and X have the same homotopy type, and the

injection of A into X and the retraction of X onto A are homotopy equivalences. A contractible space has the same homotopy type as a point. Spaces having the same homotopy type have isomorphic homotopy groups and \dagger (co)homology groups. Since the mapping cylinder $Z_f = Y \cup_f (X \times I)$ of $f \in Y^X$ contains Y as its deformation retract, it has the same homotopy type as Y . By this homotopy equivalence, f can be replaced by the injection of $X \times 1$ into Z_f . If to each topological space there corresponds a value (which may be some element of \mathbf{R} or some algebraic structure) and the values are the same for homotopy equivalent spaces, then the value is called a **homotopy type invariant**. A homotopy type invariant is a \dagger topological invariant; for example, $\pi(X; Y)$ is a homotopy type invariant of X . If a continuous mapping $f: X \rightarrow Y$ induces isomorphisms of the homotopy groups of each \dagger arcwise connected component, then f is called a **weak homotopy equivalence**. Conversely, if X and Y are CW complexes, then a weak homotopy equivalence is a homotopy equivalence (J. H. C. Whitehead).

Now we consider the category of pointed topological spaces. Let A and B be pointed topological spaces. Then the \dagger direct sum in this category is the **one-point union** (or **bouquet**) $A \vee B$ obtained from the disjoint union $A \cup B$ by identifying two base points $*_A$ and $*_B$. $A \vee B$ is identified with the subspace $(A \times *_B) \cup (*_A \times B)$ in $A \times B$. The **reduced join** (or **smash product**) of A, B is the space obtained from $A \times B$ by smashing its subspace $A \vee B$ to a point and is denoted by $A \wedge B$. We call $A \wedge S^1$ the **(reduced) suspension** of A and denote it by SA . Repeating the suspension n times, we have the **n -fold reduced suspension** of A . We call $CA = A \wedge I$ ($I = [0, 1]$) the **reduced cone** of A (I has the base point 1). For a continuous mapping $f: X \rightarrow Y$, the space obtained by identifying each point $(x, 0)$ of the base of CX with $f(x) \in Y$ is called the **reduced mapping cone** and is denoted by $C_f = Y \cup_f CX$. The **reduced join** (or **smash product**) of mappings $f: Y \rightarrow X$ and $f': Y' \rightarrow X'$ is the mapping $f \wedge f': Y \wedge Y' \rightarrow X \wedge X'$ induced from the product mapping $f \times f': Y \times Y' \rightarrow X \times X'$. The reduced join of $f: Y \rightarrow X$ and $1: S^1 \rightarrow S^1$ (identity mapping) is written as $Sf = f \wedge 1$ and is called the **suspension** of f .

G. Puppe Exact Sequences

For $f: X \rightarrow Y$ and $g: Y \rightarrow Z$, we have $g \circ f \simeq 0$ if and only if g can be extended to a continuous mapping from C_f into Z . In other words, the sequence

$$\pi(C_f; Z)_0 \xrightarrow{i^*} \pi(Y; Z)_0 \xrightarrow{f^*} \pi(X; Z)_0$$

is exact (i.e., $\text{Im } i^* = \text{Ker } f^* = f^{*-1}(0)$, where $i: Y \rightarrow C_f$ is the canonical inclusion and 0 is the class of the constant mapping). The inclusion $i: Y \rightarrow C_f$ gives rise to the reduced mapping cone C_i . We also have the canonical inclusion $i': C_f \rightarrow C_i$. Adding the term $\pi(C_i; Z)_0 \xrightarrow{i'^*}$ to the left-hand side of the sequence above, we have a new exact sequence. Continuing this process, we obtain an exact sequence of infinite length. If X, Y satisfy a suitable condition (e.g., X, Y are CW complexes), then C_i has the same homotopy type as the reduced suspension SX of X ; i'^* is equivalent to $p^*: \pi(SX; Z)_0 \rightarrow \pi(C_f; Z)_0$ induced by a mapping $p: C_f \rightarrow SX$ smashing Y to a point; and furthermore C_p has the same homotopy type as SY , and the inclusion $i_0: SX \rightarrow C_p$ is equivalent to the suspension $Sf: SX \rightarrow SY$ of f . Thus the following **Puppe exact sequence** is obtained:

$$\begin{aligned} \dots \xrightarrow{Sp^*} \pi(SC_f; Z)_0 \xrightarrow{Si^*} \pi(SY; Z)_0 \xrightarrow{Sf^*} \pi(SX; Z)_0 \\ \xrightarrow{p^*} \pi(C_f; Z)_0 \xrightarrow{i'^*} \pi(Y; Z)_0 \xrightarrow{f^*} \pi(X; Z)_0. \end{aligned}$$

In this exact sequence, if Y is a CW complex, X is a subcomplex of Y , and f is the inclusion $i: X \rightarrow Y$, then $C_i = C_f = Y \cup C_X$ is homotopy equivalent to the space $C_i/CX = Y/X$ obtained by smashing CX to a point, and an exact sequence of the following type is obtained:

$$\begin{aligned} \dots \xrightarrow{Si^*} \pi(SX; Z)_0 \xrightarrow{\delta^*} \pi(Y/X; Z)_0 \xrightarrow{p^*} \pi(Y; Z)_0 \\ \xrightarrow{i'^*} \pi(X; Z)_0. \end{aligned}$$

A sequence equivalent to the sequence $X \xrightarrow{f} Y \xrightarrow{i} C_f$ is called a **cofiber**, for which a similar exact sequence is obtained. For a continuous mapping $f: X \rightarrow Y$, consider the subspace $E_f = \{(x, \varphi) \mid f(x) = \varphi(0)\}$ of the product space $X \times Y^I$. By identifying X with $\{(x, \varphi_x) \mid \varphi_x(I) = f(x)\}$, we can regard X as a deformation retract of E_f . By putting $p_1(x, \varphi) = \varphi(1)$, we obtain a \dagger fiber space (E_f, p_1, Y) . The fiber $T_f = p_1^{-1}(*)$ is called a **mapping track** of f . Using the \dagger covering homotopy property, we see that the sequence

$$\pi(W; T_f)_0 \xrightarrow{p_*} \pi(W; X)_0 \xrightarrow{f_*} \pi(W; Y)_0$$

is exact, where $p(x, \varphi) = x$. This sequence is also extended infinitely to the left as

$$\dots \rightarrow \pi(W; \Omega X)_0 \xrightarrow{(\Omega f)_*} \pi(W; \Omega Y)_0 \xrightarrow{i_*} \pi(W; T_f)_0 \xrightarrow{p_*},$$

where i is the inclusion of the loop space ΩY into T_f and $\Omega f: \Omega X \rightarrow \Omega Y$ is the correspondence of the loops induced from f .

H. Homotopy Sets that Form Groups

If $X = SX'$ or $Y = \Omega Y'$ (or, generally, if Y is a \dagger homotopy associative $\dagger H$ -space having a \dagger homotopy inverse), then $\pi(X; Y)_0$ forms a group. In the general case the product of the

loops induces the product of $\pi(X; \Omega Y)_0$. We represent a point of SX by (x, t) ($x \in X, t \in I$) and define the mapping $\Omega_0 g: X \rightarrow \Omega Y$ for each $g: SX \rightarrow Y$ by $\Omega_0 g(x)(t) = g(x, t)$. Hence an isomorphism $\Omega_0: \pi(SX; Y)_0 \cong \pi(X; \Omega Y)_0$ is obtained. Each of the following pairs of homomorphisms is equivalent: $f_*: \pi(SX; Y)_0 \rightarrow \pi(SX; Y')_0$ and $\Omega f_*: \pi(X; \Omega Y)_0 \rightarrow \pi(X; \Omega Y')_0$; and $h^*: \pi(X'; \Omega Y)_0 \rightarrow \pi(X; \Omega Y)_0$ and $Sh^*: \pi(SX'; Y)_0 \rightarrow \pi(SX; Y)_0$.

If S^n is an n -dimensional sphere, then $\pi_n(X) = \pi(S^n; X)_0$ is the n -dimensional homotopy group (\rightarrow Section J). Let $\pi^n(X) = \pi(X; S^n)_0$. If X is a CW complex of dimension less than $2n - 1$, $\pi^n(X)$ is the cohomotopy group isomorphic to $\pi(X; \Omega S^{n+1})_0$ (\rightarrow Section I). Let K_n be an Eilenberg-MacLane space of type (\mathbb{Z}, n) . Then we have $K_n = \Omega K_{n+1}$, and if (X, A) is a pair of CW complexes, then $\pi(X/A; K_n)_0$ coincides with the cohomology group $H^n(X, A; \mathbb{Z})$. For the classifying space $B_O(B_U)$ of the infinite orthogonal group O (infinite unitary group U) (\rightarrow Section V), $\pi(X/A; B_O)$ ($\pi(X/A; B_U)$) may be considered the KO -group $KO(X, A)$ (K -group $K(X, A)$) (\rightarrow 237 K-Theory).

I. Cohomotopy Groups

K. Borsuk defined a sum of mapping classes of X into S^n (1936), which was named Borsuk's **cohomotopy group** by E. Spanier. Spanier also studied the duality of the cohomotopy group with the homotopy group and its relations to the usual cohomology groups. A cohomotopy group of (X, A) is defined to be $\pi^n(X, A) = \pi(X, A; S^n, *)$, which forms a group if $\dim X/A < 2n - 1$. A mapping $F: X/A \rightarrow S^n \times S^n$ given by $F(x) = (f(x), g(x))$ with $f, g: X/A \rightarrow S^n$ is homotopic to a mapping into $S^n \vee S^n$. If we compose F with a folding mapping of $S^n \vee S^n$ onto S^n , we obtain a mapping that represents the sum $[f] + [g]$. With each homotopy class of a continuous mapping f of an n -dimensional polyhedron K^n into an n -dimensional sphere S^n , we associate the image $f^*(u)$ of the fundamental class $u \in H^n(S^n; \mathbb{Z})$ under the induced homomorphism $f^*: H^n(S^n; \mathbb{Z}) \rightarrow H^n(K^n; \mathbb{Z})$. We then obtain a bijective relation $\pi^n(K^n) \rightarrow H^n(K^n; \mathbb{Z})$, called **Hopf's classification theorem**.

J. Homotopy Groups

Let X be a topological space with a base point $*$, $I^n = \{t = (t_1, t_2, \dots, t_n) \mid 0 \leq t_1, t_2, \dots, t_n \leq 1\}$ be the unit n -cube, and I^n its boundary. Write $\Omega^n(X, *) = X^{I^n}(I^n, *)$ (in particular, $\Omega^1(X, *)$ is the loop space), and denote by $\pi_n(X, *)$ or simply $\pi_n(X)$ the set of arcwise connected components of $\Omega^n(X, *)$, i.e., the homotopy

classes $[f]$. Using the notation of homotopy sets, we have $\pi_n(X, *) = \pi(I^n, I^n; X, *)$. If we choose the constant mapping as the base point $*$ of $\Omega^n(X, *)$, then $\Omega^n(\Omega^n(X, *), *) = \Omega^{n+n}(X, *)$. Thus $\pi_m(\Omega^n(X, *), *) = \pi_{m+n}(X, *)$. Since π_1 is the fundamental group, $\pi_n(X, *) = \pi_1(\Omega^{n-1}(X, *))$ is also a group, called the n -dimensional **homotopy group** of X with base point $*$. "Multiplication" in homotopy groups is defined as follows: Given $f_1, f_2 \in \Omega^n(X, *)$ we define $f_1 + f_2 \in \Omega^n(X, *)$ by

$$(f_1 + f_2)(t) = \begin{cases} f_1(2t_1, t_2, \dots, t_n), & 0 \leq t_1 \leq \frac{1}{2}, \\ f_2(2t_1 - 1, t_2, \dots, t_n), & \frac{1}{2} \leq t_1 \leq 1 \end{cases}$$

(Fig. 1). Then the product or sum of $[f_1]$ and $[f_2]$ is given by $[f_1 + f_2]$. The identity is the class of the constant mapping (denoted by 0), and the inverse of $[f]$ is $[\bar{f}]$, represented by $\bar{f}(t) = f(1 - t_1, t_2, \dots, t_n)$. The space $\Omega^n(X, *)$ is an H -space, where multiplication is given by the correspondence $(f_1, f_2) \rightarrow f_1 + f_2$. Since the fundamental group of an H -space is commutative, $\pi_n(X, *)$ is an Abelian group for $n \geq 2$.

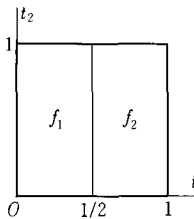


Fig. 1

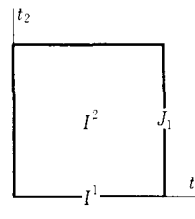


Fig. 2

Let $S^n = \{t = (t_1, \dots, t_{n+1}) \mid \sum t_i^2 = 1\}$ be the n -sphere, and take $*$ = $(1, 0, \dots, 0)$ as its base point. Suppose that we are given a continuous mapping $\psi_n: (I^n, I^n) \rightarrow (S^n, *)$ such that $\psi_n: I^n - I^n \rightarrow S^n - *$ is homeomorphic. Then the correspondence $\psi_n^*: \pi(S^n; X)_0 \rightarrow \pi_n(X, *)$ determined by $\psi_n^*[g] = [g \circ \psi_n]$ is bijective. Thus we can identify the homotopy group $\pi_n(X, *)$ with $\pi(S^n; X)_0$.

K. Relative Homotopy Groups

Suppose that we are given a topological space X and a subspace A of X sharing the same base point $*$. Identify I^{n-1} with the face $t_n = 0$ of I^n , and let J^{n-1} be the closure of $I^n - I^{n-1}$ (Fig. 2). Denote by $\pi_n(X, A, *)$ the set of homotopy classes of continuous mappings $f: (I^n, I^n, J^{n-1}) \rightarrow (X, A, *)$. Let $\Omega^n(X, A, *)$ be the mapping space consisting of such mappings f , and let $\pi_n(X, A, *) = \pi_0(\Omega^n(X, A, *))$. Since $\Omega^n(\Omega^n(X, A, *), *)$ is homeomorphic to $\Omega^{n+n}(X, A, *)$, we have $\pi_m(\Omega^n(X, A, *), *) \cong \pi_{m+n}(X, A, *)$. Thus $\pi_n(X, A, *)$ is a group for $n \geq 2$ and an Abelian group for $n \geq 3$. This group is called the n -dimensional **relative**

homotopy group of (X, A) with respect to the base point $*$, or simply the n -dimensional homotopy group of (X, A) . In the same manner as in Section J multiplication in this group can be defined using $f_1 + f_2$. Since $\Omega^n(X, *, *)$ and $\Omega^n(X, *)$ are identical, we have $\pi_n(X, *, *) = \pi_n(X, *)$. Hence homotopy groups are special cases of relative homotopy groups.

Let $g: (X, A, *) \rightarrow (Y, B, *)$ be a continuous mapping. Then a correspondence $g_*: \pi_n(X, A, *) \rightarrow \pi_n(Y, B, *)$ is obtained by $g_*[f] = [g \circ f]$, with g_* a homomorphism of homotopy groups for $n \geq 2$ and for $n = 1, A = *$. We call g_* the **homomorphism induced by g** . Let $E^n = \{t = (t_1, \dots, t_n) \mid \sum t_i^2 = 1\}$ be the unit n -cell with boundary S^{n-1} . Utilizing a suitable relative homeomorphism $\psi'_n: (I^n, J^{n-1}) \rightarrow (E^n, *)$, $\psi'_n(I^n) = S^{n-1}$, we obtain a one-to-one correspondence $\psi'^*_n: \pi(E^n, S^{n-1}; X, A)_0 \rightarrow \pi_n(X, A, *)$, and $\Omega^n(X, A, *)$ is homeomorphic (via ψ'_n) to the mapping space $X^{E^n}(S^{n-1}, *, A, *)$.

L. Homotopy Exact Sequences

Given an element $\alpha = [f] \in \pi_n(X, A, *)$, and letting $\partial\alpha = [f|I^{n-1}] \in \pi_{n-1}(A, *)$, we obtain a homomorphism ($n \geq 2$) $\partial: \pi_n(X, A, *) \rightarrow \pi_{n-1}(A, *)$, which is called the **boundary homomorphism**. Furthermore, we have the following exact sequence involving homomorphisms $i_* \cdot j_*$ induced by two inclusions $i: (A, *) \rightarrow (X, *)$, $j: (X, *, *) \rightarrow (X, A, *)$:

$$\begin{aligned} \dots \xrightarrow{\partial} \pi_n(A, *) \xrightarrow{i_*} \pi_n(X, *) \xrightarrow{j_*} \pi_n(X, A, *) \\ \xrightarrow{\partial} \dots \rightarrow \pi_1(X, *) \xrightarrow{j_*} \pi_1(X, A, *) \xrightarrow{\partial} \pi_0(A) \xrightarrow{i_*} \pi_0(X). \end{aligned}$$

This sequence is called the **homotopy exact sequence** of the pair (X, A) . A system of topological spaces $X \supset A \supset B \ni *$ is called a **triple**. In this homotopy exact sequence, if we replace $(A, *)$, $(X, *)$ by $(A, B, *)$, $(X, B, *)$, respectively, we obtain an exact sequence, called the **homotopy exact sequence of the triple** (X, A, B) .

The homotopy group $\pi_n(A \times B)$ of the product space is isomorphic to the direct sum $\pi_n(A) + \pi_n(B)$, and the projections $p(p'): A \times B \rightarrow A(B)$ of the product space induce the projections from $\pi_n(A \times B)$ onto the direct summands $\pi_n(A)$, $\pi_n(B)$. This is a special case of the Hurewicz-Steenrod isomorphism theorem in fiber spaces (\rightarrow 148 Fiber Spaces). Setting $A \vee B = (A \times *) \cup (* \times B)$, we obtain a direct sum decomposition $\pi_n(A \vee B) \cong \pi_n(A) + \pi_n(B) + \pi_{n+1}(A \times B, A \vee B)$. Next we consider a fixed pair (X, A) and move the base point $*$ to investigate its effect on the elements of the homotopy group. Suppose that we are given a path $h: I \rightarrow A$ with terminal point $* = h(1)$ and an element $\alpha \in \pi_n(X, A, *)$ ($\alpha = [f], f: (I^n, I^n, J^{n-1}) \rightarrow (X, A, *)$). By the homotopy extension property, we can construct a homotopy $f_\theta: (I^n, I^n)$

$\rightarrow (X, A)$ satisfying $f_\theta(J^{n-1}) = h(\theta)$ and $f_1 = f$. Then the homotopy class $[f_\theta]$ of f_θ with respect to the base point $*' = h(0)$ is determined only by α and the homotopy class ω of the path h . We denote the homotopy class $[f_\theta]$ by $\alpha^\omega \in \pi_n(X, A, *')$. The correspondence $\alpha \rightarrow \alpha^\omega$ is a group isomorphism, and $(\alpha^\omega)^{\omega'} = \alpha^{\omega\omega'}$. Thus if A is arcwise connected, $\pi_n(X, A, *)$ is isomorphic to $\pi_n(X, A, *')$. Hence, in this case, we may simply write $\pi_n(X, A)$ instead of $\pi_n(X, A, *)$. When $* = *'$, the correspondence $\alpha \rightarrow \alpha^\omega$ determines the action of the group $\pi_1(A, *)$ on $\pi_n(X, A, *)$. Given an element $\alpha \in \pi_n(X, *)$ and a class ω of paths in X , we define $\alpha^\omega \in \pi_n(X, *')$ as for relative homotopy. Specifically, if $\omega \in \pi_1(X, *)$, then $\alpha^\omega - \alpha$ coincides with the Whitehead product $[\omega, \alpha]$ (when $n = 1$, we have $\alpha^\omega \cdot \alpha^{-1} = [\omega, \alpha] = \omega\alpha\omega^{-1}\alpha^{-1}$) (\rightarrow Section P).

A pair (X, A) consisting of a topological space X and an arcwise connected subspace A of X is said to be **n -simple** if the operation of $\pi_1(A)$ on $\pi_n(X, A)$ is trivial. Similarly, an arcwise connected space X is called **n -simple** if the operation of $\pi_1(X)$ on $\pi_n(X)$ is trivial. For example, a pair (X, A) consisting of an H -space X and an H -subspace A is **simple**, i.e., n -simple for each n . If a topological space X satisfies $\pi_i(X) = 0$ ($0 \leq i \leq n$), then X is said to be **n -connected**. 0-connectedness coincides with arcwise connectedness and 1-connectedness means \dagger simple connectedness. S^n is $(n-1)$ -connected. A pair (X, A) is said to be **n -connected** if $\pi_0(A) = \pi_0(X) = \pi_i(X, A) = 0$ ($1 \leq i \leq n$), and (E^n, S^{n-1}) is $(n-1)$ -connected.

M. Homotopy Groups of Triads

Let $(X; A, B, *)$ be a system, called a **triad**, of a topological space X and its subspaces A, B satisfying $A \cap B \ni *$ (base point). Let $\pi_n(X; A, B, *) = \pi_{n-1}(\Omega^1(X, B), \Omega^1(A, A \cap B), *)$ ($n \geq 2$); $\pi_n(X; A, B, *)$ is a group for $n \geq 3$ and an Abelian group for $n \geq 4$. We call $\pi_n(X; A, B, *)$ the **homotopy group of the triad**. From the homotopy exact sequence of the pair, we obtain the following **homotopy exact sequence of the triad**:

$$\begin{aligned} \dots \xrightarrow{\partial} \pi_j(A, A \cap B, *) \xrightarrow{i_*} \pi_j(X, B, *) \\ \xrightarrow{j_*} \pi_j(X; A, B, *) \xrightarrow{\partial} \pi_{j-1}(A, A \cap B, *) \xrightarrow{i_*} \dots \end{aligned}$$

Assume for simplicity that $A \cap B$ is simply connected, $X = \text{Int } A \cup \text{Int } B$ ($\text{Int } A$ is the \dagger interior of A), $(A, A \cap B)$ is m -connected, and $(B, A \cap B)$ is n -connected. Then $(X; A, B)$ is $(m+n)$ -connected, i.e., $\pi_j(X; A, B, *) = 0$ ($2 \leq j \leq m+n$) (**Blakers-Massey theorem**).

Furthermore, in this case we have a replica of the \dagger excision isomorphism in homology theory for $j < m+n$; that is, we have the isomorphism $i_* \cdot \pi_j(A, A \cap B, *) \cong \pi_j(X, B, *)$ in-

duced by the inclusion $i: (A, A \cap B) \rightarrow (X, B)$. On the other hand, $\pi_{m+n+1}(X; A, B, *)$ is isomorphic to $\pi_{m+1}(A, A \cap B, *) \otimes \pi_{n+1}(B, A \cap B, *)$. This shows that the excision isomorphism does not always hold for homotopy groups, an important difference from homology theory. However, if we replace the excision axiom by the Hurewicz-Steinrod isomorphism theorem, which is valid for fiber spaces (\rightarrow 148 Fiber Spaces), then we can construct homotopy theory axiomatically in the same manner as homology theory (\rightarrow 201 Homology Theory).

N. The Hurewicz Isomorphism Theorem

The **Hurewicz homomorphism** τ of $\pi_n(X, A)$ into the n -dimensional integral homology group $H_n(X, A)$ is defined by $\tau([f]) = f_*(\varepsilon_n)$ (where ε_n is a generator of $H_n(I^n, \dot{I}^n)$). Then we have the **Hurewicz isomorphism theorem**: Suppose that the pair (X, A) is n -simple (e.g., $A = *$) and $(n-1)$ -connected. Then we have $H_i(X, A) = 0$ ($i < n$) and the isomorphism $\tau: \pi_n(X, A) \cong H_n(X, A)$ (for $n = 1 \rightarrow 170$ Fundamental Groups). Let X, Y be simply connected topological spaces, and let $f: X \rightarrow Y$ be a continuous mapping. Then the following two conditions are equivalent: (1) $f_*: \pi_i(X) \rightarrow \pi_i(Y)$ is injective for $i < n$ and surjective for $i \leq n$. (2) $f_*: H_i(X) \rightarrow H_i(Y)$ is injective for $i < n$ and surjective for $i \leq n$ (**J. H. C. Whitehead's theorem**).

J.-P. Serre generalized these theorems as follows: A family \mathcal{C} of Abelian groups satisfying condition (i) is called a **class of Abelian groups**: (i) If a sequence $F \rightarrow G \rightarrow H$ of Abelian groups is exact and $F, H \in \mathcal{C}$, then $G \in \mathcal{C}$. Furthermore, we consider the following conditions: (ii) The tensor product $G \otimes F$ of an arbitrary Abelian group F with an element $G \in \mathcal{C}$ also belongs to \mathcal{C} . (ii') If both $F, G \in \mathcal{C}$, then $F \otimes G, \text{Tor}(F, G) \in \mathcal{C}$. (iii) If $G \in \mathcal{C}$, then its i th homology group $H_i(G) \in \mathcal{C}$ ($i > 0$). Condition (ii') is implied by (ii). A homomorphism $f: F \rightarrow G$ is called \mathcal{C} -injective if $\text{Ker } f \in \mathcal{C}$, \mathcal{C} -surjective if $\text{Coker } f = G/\text{Im } f \in \mathcal{C}$, and a \mathcal{C} -isomorphism if f is \mathcal{C} -injective and \mathcal{C} -surjective. Two Abelian groups G and G' are called \mathcal{C} -isomorphic if there exist \mathcal{C} -isomorphisms $f: F \rightarrow G$ and $f': F \rightarrow G'$. In particular, if the class \mathcal{C}_0 consists of only the trivial group 0, then concepts such as \mathcal{C}_0 -isomorphism coincide with the usual concepts of isomorphism, and so on. Let \mathcal{C}_p be the class of finite Abelian groups whose orders are relatively prime to a fixed prime number p . Here, instead of the terms \mathcal{C}_p -isomorphism and so on, we use the terms **mod p isomorphism** and so on. Let \mathcal{D} be a class of finitely generated Abelian groups. Then \mathcal{C}_p satisfies conditions (ii) and (iii), and \mathcal{D} satisfies (ii') and (iii).

We have the following **generalized Hurewicz theorem**: (A) Suppose that a class \mathcal{C} satisfies (ii) and (iii) and we are given a 2-connected pair (X, A) of simply connected spaces X, A . If $\pi_i(X, A) \in \mathcal{C}$ ($i < n$), then $H_i(X, A) \in \mathcal{C}$, and $\tau: \pi_n(X, A) \rightarrow H_n(X, A)$ is a \mathcal{C} -isomorphism. (B) Suppose that $A = *$, \mathcal{C} satisfies conditions (ii') and (iii), and X is simply connected. Then an assertion similar to (A) holds. In particular, a simply connected space X having finitely generated homology groups (e.g., a simply connected finite polyhedron) has finitely generated homotopy groups. As a corollary to theorem (A), we obtain a **generalized Whitehead theorem**. In particular, applying the theorem to the class $\mathcal{D} \cap \mathcal{C}_p$, we obtain the following frequently used theorem: Suppose that we are given simply connected spaces X, Y whose homology groups are finitely generated and $f: X \rightarrow Y$ satisfies $f_*\pi_2(X) = \pi_2(Y)$. Then the following two conditions are equivalent: (1) $f_*: \pi_i(X) \rightarrow \pi_i(Y)$ is a mod p isomorphism for $i < n$ and a mod p surjection for $i = n$. (2) $f_*: H_i(X, \mathbb{Z}_p) \rightarrow H_i(Y, \mathbb{Z}_p)$ is an isomorphism for $i < n$ and a surjection for $i = n$ (where $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$). The theory above, which makes use of the notion of class \mathcal{C} , is an example of **Serre's \mathcal{C} -theory**. Concepts such as i spectral sequences for fiber spaces and i n-connective fiber spaces are important tools in Serre's \mathcal{C} -theory (\rightarrow 148 Fiber Spaces).

To calculate homotopy groups, we use notions such as exact sequences, fiber spaces, (co)homology groups of n -connective fiber spaces, and i Postnikov systems. Given an arbitrary group (more generally, a Postnikov system), there exists a CW complex having the given group (system) as its homotopy group (Postnikov system) (**realization theorem of homotopy groups**). For an arbitrary arcwise connected topological space X there exist topological spaces (X, n) and continuous mappings $p_n: (X, n+1) \rightarrow (X, n)$ ($n = 1, 2, \dots$) satisfying the following two conditions: (i) $((X, n+1), p_n, (X, n))$ is a fiber space whose fiber is an i Eilenberg-MacLane space. (ii) $(X, 1) = X$, and $((X, n+1), p_1 \circ \dots \circ p_n, X)$ is an n -connective fiber space. The method of obtaining the homotopy group $\pi_n(X) \cong H_n((X, n))$ by computing (co)homology groups of (X, n) is called a **killing method**.

O. Homotopy Operations

Let X, Y, X', Y' be topological spaces. If to each continuous mapping $f \in Y^X$ there corresponds a homotopy class $\Phi(f) \in \pi(X'; Y')$ that is a homotopy invariant of f (satisfying a certain naturality condition), then Φ is called a **homotopy operation**. More generally, we

may consider the case where Φ is a mapping from $\pi(X_1; Y_1) \times \dots \times \pi(X_r; Y_r)$ into $\pi(X'; Y')$. The **naturality** of Φ is defined as follows: Consider the †category \mathcal{C} of topological spaces (or its subcategory). Let $Y = Y'$ be an arbitrary †object of \mathcal{C} , and fix X and X' . In this case, the naturality of $\Phi_Y: \pi(X; Y) \rightarrow \pi(X'; Y)$ is defined to be the commutativity of the diagram:

$$\begin{array}{ccc} \pi(X; Y) & \xrightarrow{\Phi_Y} & \pi(X'; Y) \\ \downarrow g_* & & \downarrow g_* \\ \pi(X; Z) & \xrightarrow{\Phi_Z} & \pi(X'; Z) \end{array}$$

i.e., $g_* \circ \Phi_Y = \Phi_Z \circ g_*$ for an arbitrary †morphism (i.e., continuous mapping) $g: Y \rightarrow Z$ of the category \mathcal{C} . Similarly, when objects Y, Y' of the category \mathcal{C} are fixed and $X = X'$ is an arbitrary object of \mathcal{C} , to say that a homotopy operation $\Phi_X: \pi(X; Y) \rightarrow \pi(X; Y')$ is **natural** means that $h^* \circ \Phi_X = \Phi_{X'} \circ h^*$ for an arbitrary morphism $h: W \rightarrow X$.

We have the following theorem: In the category of topological spaces and continuous mappings, the homotopy operations $\Phi_Y: \pi(X; Y) \rightarrow \pi(X'; Y)$ and the elements of $\pi(X'; X)$ are in one-to-one correspondence. The correspondence is obtained by associating a homotopy operation $\Phi(\beta) = \beta \circ \alpha$ ($\beta \in \pi(X; Y)$) with each $\alpha \in \pi(X'; X)$. Similarly, the homotopy operations $\Phi_X: \pi(X; Y) \rightarrow \pi(X; Y')$ and the elements of $\pi(Y; Y')$ are in one-to-one correspondence. This theorem holds also for the case involving several variables if we consider $\pi(X'; X_1 \vee X_2 \vee \dots)$ or $\pi(Y_1 \times Y_2 \times \dots; Y')$ instead of $\pi(X'; Y)$ or $\pi(X; Y')$. The theorem remains valid if we replace the spaces X, Y by systems of spaces.

P. Homotopy Operations in Homotopy Groups

(1) If X, X' are spheres S^n, S^p with base points and Y, Y' are topological spaces with base points, a homotopy operation $\Phi_Y: \pi_n(Y) \rightarrow \pi_p(Y)$ is said to be of type (n, p) . By the theorem in Section O, the homotopy operations of type (n, p) are in one-to-one correspondence with the elements of the homotopy group of the sphere $\pi_p(S^n)$.

(2) As an example of the 2-variable homotopy operations $\Phi_Y: \pi_m(Y) \times \pi_n(Y) \rightarrow \pi_p(Y)$ of type $(m, n; p)$ we have the Whitehead product defined as follows: Suppose that $\alpha \in \pi_m(Y)$, $\beta \in \pi_n(Y)$ are elements represented by $f: (I^m, i^m) \rightarrow (Y, *)$ and $g: (I^n, i^n) \rightarrow (Y, *)$, respectively. Define a continuous mapping F from the boundary $I^{m+n} = (I^m \times I^n) \cup (i^m \times I^n) \cup (I^m \times i^n)$ into Y by $F(x, y) = f(x)$ for $(x, y) \in I^m \times I^n$ and $F(x, y) = g(y)$ for $(x, y) \in i^m \times I^n$. Since I^{m+n} is homeomorphic to S^{m+n-1} , we can identify them. The homotopy class represented by F is an element of $\pi_{m+n-1}(Y)$ determined by α

and β , denoted by $[\alpha, \beta]$ and called the **Whitehead product** of α and β (J. H. C. Whitehead, *Ann. Math.*, (2) 42 (1941)). The Whitehead product is a homotopy operation of type $(m, n; m+n-1)$. Let $\psi_m: (I^m, i^m) \rightarrow (S^m, *)$ be a mapping that smashes i^m to a point. The product of ψ_m and ψ_n defines a mapping $\psi_{m,n}: S^{m+n-1} \rightarrow S^m \vee S^n = (S^m \times *) \cup (* \times S^n)$. Let $i \in \pi_m(S^m \vee S^n)$, $i' \in \pi_n(S^m \vee S^n)$ be the homotopy classes of the natural inclusions of S^m, S^n into $S^m \vee S^n$; then the homotopy class of $\psi_{m,n}$ is $[i, i']$. G. W. Whitehead showed that a direct sum decomposition $\pi_p(S^m \vee S^n) = i_* \pi_p(S^m) + i'_* \pi_p(S^n) + [i, i']_* \pi_p(S^{m+n-1})$ ($i_*, i'_*, [i, i']_*$ are injective) holds for $1 < p < m+n+\min(m, n)-3$. Furthermore, P. J. Hilton showed that for general $p > 1$, $\pi_p(S^m \vee S^n)$ is the direct sum of the images of injections $i_*, i'_*, [i, i']_*, [[i, i'], i]_*, [[i, i'], i']_*,$ etc. The homotopy operations of type $(m, n; p)$ are in one-to-one correspondence with the elements of $\pi_p(S^m \vee S^n)$; hence such operations can be constructed by means of composition and the Whitehead product. The last proposition is also valid for homotopy operations of type $(m_1, \dots, m_r; p)$. The Whitehead product $[\alpha, \beta]$ ($\alpha \in \pi_m(X)$, $\beta \in \pi_n(X)$) is distributive with respect to α (resp. β) for $m > 1$ ($n > 1$), and we have $[\beta, \alpha] = (-1)^{mn} [\alpha, \beta]$ and $f_* [\alpha, \beta] = [f_* \alpha, f_* \beta]$ for $f: X \rightarrow Y$. Moreover, for $\gamma \in \pi_r(X)$ the **Jacobi identity** holds: $(-1)^{mr} [[\alpha, \beta], \gamma] + (-1)^{mn} [[\beta, \gamma], \alpha] + (-1)^{rn} [[\gamma, \alpha], \beta] = 0$ (M. Nakaoka and H. Toda; H. Uehara and W. S. Massey; Hilton).

Q. Suspensions and Generalized Hopf Invariants

We denote by $\alpha \wedge \beta \in \pi(X \wedge X'; Y \wedge Y')_0$ the class of the reduced join of f, g , where f represents $\alpha \in \pi(X; Y)_0$ and g represents $\beta \in \pi(X'; Y')_0$. We call $\alpha \wedge \beta$ the **reduced join** of α and β . In particular, if $Y = Y' = S^1$, β is the identity mapping of S^1 , and α is represented by f , then $\alpha \wedge \beta$ is called the **suspension** of α and is denoted by $S\alpha$. $S\alpha$ is the class of the suspension Sf of f and belongs to $\pi(SX; SY)_0$, where SX indicates the reduced suspension of X . The suspension $S\alpha$ is often denoted by $E\alpha$ in reference to the German term *Einhängung*. The identity mapping 1 of SY gives rise to an injection $i = \Omega_0 1$ sending Y into the loop space $\Omega(SY)$ determined by the formula $i(y)(t) = (y, t)$. Then we have

$$i_* = \Omega_0 \circ S: \pi(X; Y)_0 \rightarrow \pi(SX; SY)_0 \cong \pi(X; \Omega SY)_0,$$

and S and i_* are equivalent. Let Y^k be the identifying space Y^k / \sim , where Y^k is the product space $Y \times \dots \times Y$ of k copies of Y and \sim

is the equivalence relation determined by

$$(*, y_1, y_2, \dots, y_{k-1}) \sim (y_1, *, y_2, \dots, y_{k-1}) \sim \dots \sim (y_1, \dots, y_{k-1}, *).$$

Denote by $Y_\infty = \bigcup_k Y_k$ the limit space with respect to the injection $Y_{k-1} \rightarrow Y_k$ given by $(y_1, \dots, y_{k-1}) \rightarrow (y_1, \dots, y_{k-1}, *)$ and call it the **reduced product space** of Y . Let Y be a CW complex of 0-section $*$. The mapping $i: Y = Y_1 \rightarrow \Omega SY$ can then be extended to $\bar{i}: Y_\infty \rightarrow \Omega SY$, where \bar{i} is a weak homotopy equivalence. If X is also a CW complex, then $\Omega_0^{-1} \circ i_*: \pi(X; Y_\infty)_0 \rightarrow \pi(SX; SY)_0$ is bijective. By smashing the subset Y of Y_2 , we have $Y \wedge Y = Y_2/Y$. This smashing mapping can be extended to $h: Y_\infty \rightarrow (Y \wedge Y)_\infty$ (I. M. James). Utilizing $h_*: \pi(X; Y_\infty)_0 \rightarrow \pi(X; (Y \wedge Y)_\infty)_0$ and the bijection $\Omega_0^{-1} \circ \bar{i}_*$, we obtain a correspondence $H: \pi(SX; SY)_0 \rightarrow \pi(SX; S(Y \wedge Y)_0)$. We call $H(\alpha)$ the **generalized Hopf invariant** of α . When $X = S^{2n-2}$, $Y = S^{n-1}$, H is equivalent to the Hopf invariant $\gamma: \pi_{2n-1}(S^n) \rightarrow \mathbb{Z}$ (\rightarrow Section U). In general, we have $H \circ S = 0$, and the exactness of $\xrightarrow{S} \xrightarrow{H}$ holds under various conditions.

Denote also by \circ the composition of homotopy classes; then we have $S(\alpha \circ \beta) = S\alpha \circ S\beta$ and $H(\alpha \circ S\beta) = H\alpha \circ S\beta$. Also, $H(S\alpha \circ \beta) = S(\alpha \wedge \alpha) \circ H\beta$. Under the condition $i < 3n - 3$, we have $(\alpha_1 + \alpha_2) \circ \beta = \alpha_1 \circ \beta + \alpha_2 \circ \beta + [\alpha_1, \alpha_2] \circ H(\beta)$ for $\alpha_1, \alpha_2 \in \pi_n(X)$ and $\beta \in \pi_i(S^n)$ (G. W. Whitehead). Thus the composition $\alpha \circ \beta$ is not always left distributive but is always right distributive, and $\alpha \circ \beta$ is left distributive if $\beta = S\beta'$. The composition is defined over the stable homotopy groups G_r of spheres (\rightarrow Section U): $\alpha \circ \beta \in G_{p+q}$ ($\alpha \in G_p, \beta \in G_q$). It is distributive and satisfies $\beta \circ \alpha = (-1)^{pq} \alpha \circ \beta$.

When Y and Y' are Eilenberg-MacLane spaces, B_O , and B_U , we have t cohomology operations on cohomology groups $H^n(-; \mathbb{I})$, KO groups, and K groups, respectively. As typical examples there are t Steenrod square operations $Sq^i: H^n(X; \mathbb{Z}_2) \rightarrow H^{n+i}(X; \mathbb{Z}_2)$, t Steenrod p th power operations $\mathcal{P}^i: H^n(X; \mathbb{Z}_p) \rightarrow H^{n+2i(p-1)}(X; \mathbb{Z}_p)$, t Chern characters $ch^n: K(X) \rightarrow H^{2n}(X; \mathbb{Q})$ (\mathbb{Q} : rational field), t Adams operations $\Phi_i: KO(X) \rightarrow KO(X)$ ($K(X) \rightarrow K(X)$). They are all homomorphisms (\rightarrow 64 Cohomology Operations; 237 K-Theory).

R. Secondary Compositions

Suppose that $\alpha \circ \beta = 0$, $\beta \circ \gamma = 0$ for $\gamma \in \pi(W; X)_0$, $\beta \in \pi(X; Y)_0$, $\alpha \in \pi(Y; Z)_0$. In the commutative diagram of Puppe exact sequences

$$\begin{array}{ccccccc} S\gamma^* & & & & & & \\ \downarrow \downarrow_* & \xrightarrow{p^*} & \downarrow \downarrow_* & \xrightarrow{i^*} & \downarrow \downarrow_* & \xrightarrow{\gamma^*} & \\ \pi(SW; Y)_0 & \rightarrow & \pi(C_\gamma; Y)_0 & \rightarrow & \pi(X; Y)_0 & \rightarrow & \\ S\gamma^* & & & & & & \\ \downarrow \downarrow_* & \xrightarrow{p^*} & \downarrow \downarrow_* & \xrightarrow{i^*} & \downarrow \downarrow_* & \xrightarrow{\gamma^*} & \\ \pi(SW; Z)_0 & \rightarrow & \pi(C_\gamma; Z)_0 & \rightarrow & \pi(X; Z)_0 & \rightarrow & \end{array}$$

the set of elements $\bar{\beta}$ in $\pi(SW; Z)$ such that $p^*(\bar{\beta}) \in \alpha_* i^{*-1}(\beta)$ is denoted by $\{\alpha, \beta, \gamma\}$ and is called a **secondary composition** or **Toda bracket**. If θ, η are elements of $\pi(SW; Y)_0$, $\pi(SX; Z)_0$, respectively, then we have $\{\alpha, \beta, \gamma\} + \alpha_* \theta = \{\alpha, \beta, \gamma\}$, $\{\alpha, \beta, \gamma\} + S\gamma^* \eta = \{\alpha, \beta, \gamma\}$. Hence we may consider the set $\{\alpha, \beta, \gamma\}$ to be a residue class modulo a submodule generated by $\alpha_* \pi(SW; Y)_0$ and $S\gamma^* \pi(SX; Z)_0$.

The secondary composition $\{\alpha, \beta, \gamma\}$ has the following properties: (i) $\{\alpha, \beta, \gamma\}$ is linear with respect to α, β, γ (if the sum is defined); (ii) $\alpha \circ \{\beta, \gamma, \delta\} = \{\alpha, \beta, \gamma\} \circ (-S\delta)$; (iii) $S\{\alpha, \beta, \gamma\} \equiv -\{S\alpha, S\beta, S\gamma\}$; (iv) $\alpha \circ \{\beta, \gamma, \delta\} \equiv \{\alpha \circ \beta, \gamma, \delta\}$, $\{\alpha \circ \beta, \gamma, \delta\} \equiv \{\alpha, \beta \circ \gamma, \delta\}, \dots$; (v) $\{\{\alpha, \beta, \gamma\}, S\delta, S\epsilon\} + \{\alpha, \{\beta, \gamma, \delta\}, S\epsilon\} + \{\alpha, \beta, \{\gamma, \delta, \epsilon\}\} \equiv 0$. Suppose that the spaces X, Y, Z, W are spheres. Then by (iii) the secondary composition $\{\alpha, \beta, \gamma\} \in G_{p+q+r+1}/(\alpha \circ G_{q+r+1} + \gamma \circ G_{p+q+1})$ is defined in the stable homotopy groups $G_r = \lim_{n \rightarrow \infty} \pi_{n+r}(S^n)$ of spheres. From this we obtain (vi) $\{\gamma, \beta, \alpha\} = (-1)^{pq+qr+rp+1} \{\alpha, \beta, \gamma\}$ and (vii) $(-1)^{pr} \{\alpha, \beta, \gamma\} + (-1)^{qp} \{\beta, \gamma, \alpha\} + (-1)^{rq} \{\gamma, \alpha, \beta\} \equiv 0$.

S. Functional Operations

Let Φ be an operation corresponding to α and γ be the class of f . We put $\Phi_f(\beta) = \{\alpha, \beta, \gamma\}$ and call Φ_f a **functional Φ -operation**. When Φ is a cohomology operation, Φ_f is called a **functional cohomology operation**. Then $\Phi_f(\beta)$ is defined for β satisfying $f^*(\beta) = \Phi(\beta) = 0$, and $\Phi_f(\beta)$ is determined modulo $\text{Im } Sf^* + \text{Im } \Phi$. For $f: S^{n+k} \rightarrow S^n$, $k = 2i(p-1) - 1$, we denote by $H_p(f) \cdot e_{n+k+1} \in H^{n+k+1}(S^{n+k+1}; \mathbb{Z}_p)$ the image of a generator e_n of $H^n(S^n; \mathbb{Z}_p)$ under the functional \mathcal{P}_f^i operation. Then the **Hopf invariant modulo p** (or **mod p Hopf invariant**) $H_p: \pi_{n+k}(S^n) \rightarrow \mathbb{Z}_p$ is obtained (we use Sq^{2i} for $p = 2$). The following statements are equivalent: (i) The mod 2 Hopf invariant is not trivial ($H_2 \neq 0$); (ii) there exists a mapping: $S^{2k+1} \rightarrow S^{k+1}$ of Hopf invariant 1; (iii) S^k is an H -space; (iv) the Whitehead product $[i, i]$ of a generator i of $\pi_k(S^k)$ vanishes. Also, $H_2 \neq 0$ if and only if $k = 2, 4, 8$ (J. Adams), and for an odd prime p , $H_p \neq 0$ if and only if $k = 2p - 3$ (A. L. Liulevicius; N. Shimada and T. Yamanoshita).

T. Stable Homotopy Groups and Spectra

The homotopy set $\pi(S^n X, S^n Y)_0$ for n -fold iterated suspensions $S^n X = X \wedge S^n = SS^{n-1} X$ and $S^n Y$, forms a group (an Abelian group) if $n \geq 1$ ($n \geq 2$). The limit $\pi^*(X; Y) = \lim \pi(S^n X; S^n Y)_0$ with respect to the suspension homomorphisms $S: \pi(S^n X; S^n Y)_0 \rightarrow \pi(S^{n+1} X; S^{n+1} Y)_0$ is called a **stable homotopy group** of X and

Y . For an r -connected space Y and a CW-complex X , $S: \pi(X; Y)_0 \rightarrow \pi(SX; SY)_0$ is bijective if $\dim X \leq 2r$ and surjective if $\dim X \leq 2r + 1$ (**generalized suspension theorem**). Thus, if X is a finite-dimensional CW-complex, $\pi^s(X, Y)$ is isomorphic to $\pi(S^n X, S^n Y)_0$ for sufficiently large n . To discuss stable homotopy groups more generally, the following concept of spectra is used. A system $E = \{E_k, \varepsilon_k\}$ which consists of CW-complexes E_k and continuous mappings $\varepsilon_k: SE_k \rightarrow E_{k+1}$ is called a **spectrum**. When $E_k = S^k$ and $\varepsilon_k = 1_{k+1}: S^k \rightarrow S^{k+1}$, $S = \{S^k, 1_{k+1}\}$ is called a **sphere spectrum**. When $E_k = K(G, k)$ (\dagger Eilenberg-MacLane complex) and ε_k induces a homotopy equivalence $K(G, k) \simeq \Omega K(G, k + 1)$, $HG = \{K(G, k), \varepsilon_k\}$ is called an **Eilenberg-MacLane spectrum**. As in the latter, a spectrum E in which ε_k induces a homotopy equivalence $E_k \simeq \Omega E_{k+1}$ is called an Ω -spectrum. By Bott's periodicity, Ω -spectra $KU = \{Z \times B_\infty, U, Z \times B_\infty, U, \dots\}$ and $KO = \{Z \times B_\infty, U/O, Sp/U, Sp, Z \times B_\infty, U/Sp, SO/U, O, Z \times B_\infty, \dots\}$ are obtained (\rightarrow Section V). Also, using \dagger Thom complexes, **Thom spectra** MU, MO , etc. are obtained. Given a spectrum E , by putting $E^n(X, A) = \lim \pi(S^k(X/A); E_{k+1})_0$ for each pair (X, A) of CW-complexes, we obtain a **generalized cohomology theory with E-coefficient**; and by putting $E_n(X, A) = \lim \pi_k(E_{k-n} \wedge (X/A))$, we obtain a **generalized homology theory with E-coefficient**. \dagger Generalized (co)homology theory on (finite) CW-complexes can be represented by a suitable spectrum (G. W. Whitehead, E. H. Brown, Adams). Corresponding to $E = S, HG, KU, MU$, etc., we have stable (co)homotopy groups, G -coefficient (co)homology groups, K -groups, \dagger (co)bordism groups, etc., respectively (\rightarrow 201 Homology Theory).

U. Homotopy Groups of Spheres

The spheres S^n and their homotopy groups are basic objects in homotopy theory. Although much research has been done concerning these objects, there are still open problems.

S^n is $(n - 1)$ -connected: $\pi_i(S^n) = 0$ ($i < n$). The fact that $\pi_n(S^n) \cong \mathbb{Z}$ (infinite cyclic group) was obtained from the \dagger Brouwer mapping theorem. Also, $\pi_i(S^1) = 0$ ($i > 1$) follows from the fact that the \dagger universal covering space of S^1 is contractible. Suppose that we are given a continuous mapping $f: S^{2n-1} \rightarrow S^n$. We approximate it by a \dagger simplicial mapping φ . Then the inverse image $\varphi^{-1}(\ast)$ of a point \ast in the interior of an n -simplex of S^n is an $(n - 1)$ -dimensional \dagger pseudomanifold which is orientable by means of a suitable generator $\varepsilon \in H_{n-1}(\varphi^{-1}(\ast))$. The boundary isomorphism $\partial: H_n(S^{2n-1}, \varphi^{-1}(\ast)) \cong H_{n-1}(\varphi^{-1}(\ast))$ and the homomorphism $\varphi_\ast:$

$H_n(S^{2n-1}, \varphi^{-1}(\ast)) \rightarrow H_n(S^n, \ast)$ give rise to an integer $\gamma(\varphi)$ determined by the relation $\varphi_\ast \partial^{-1}(\varepsilon) = \gamma(\varphi) \varepsilon_n$ (ε_n is an orientation of S^n). This integer is independent of the choice of φ , so we can set $\gamma(f) = \gamma(\varphi)$. Then $f \simeq g$ implies that $\gamma(f) = \gamma(g)$. We call $\gamma(f)$ the **Hopf invariant** of f . H. Hopf defined γ and showed $\gamma: \pi_3(S^2) \cong \mathbb{Z}$ (1931); $\gamma(\pi_{2n-1}(S^n)) = 0$ for odd n ; $\gamma(\pi_{2n-1}(S^n)) \subset 2\mathbb{Z}$ for even n ; and $\gamma(\pi_{2n-1}(S^n)) = \mathbb{Z}$ for $n = 4, 8$ (1935). H. Freudenthal defined a homomorphism $E: \pi_i(S^n) \rightarrow \pi_{i+1}(S^{n+1})$, $E[f] = [Sf]$, and proved the **Freudenthal theorem**: (1) E is an isomorphism for $i < 2n - 1$; (2) E is a surjection for $i = 2n - 1$; and (3) the image of E coincides with the kernel of γ for $i = 2n$. Furthermore he obtained $\pi_{n+1}(S^n) \cong \mathbb{Z}_2$ ($n \geq 3$) (1937). For $n = 2, 4, 8$, a mapping $f: S^{2n-1} \rightarrow S^n$ (\dagger Hopf mapping) such that $\gamma(f) = 1$ (given by Hopf) is the projection of a \dagger fiber bundle S^{2n-1} over the base space S^n , and the correspondence $(\alpha, \beta) \rightarrow E\alpha + f_\ast \beta$ gives an isomorphism $\pi_{i-1}(S^{n-1}) + \pi_i(S^{2n-1})$ (direct sum) $\cong \pi_i(S^n)$. Hence we obtain $\pi_4(S^2) = \mathbb{Z}_2$. It was shown by G. W. Whitehead and L. S. Pontryagin that $\pi_{n+2}(S^n)$ ($n \geq 3$) is isomorphic to \mathbb{Z}_2 (1949). Whitehead also defined a **generalized Hopf homomorphism** $H: \pi_i(S^n) \rightarrow \pi_i(S^{2n-1})$ for a range of $i < 3n - 3$, and this restriction on the dimension was removed by P. J. Hilton and I. M. James. Using H , many nontrivial results concerning $\pi_i(S^n)$ have been obtained. Serre obtained the following (1951–1953): $\pi_i(S^n)$ is finite except when $i = n$ or $i = 4m - 1$ and $n = 2m$. Furthermore, $\pi_{4m-1}(S^{2m})$ is the direct sum of \mathbb{Z} and a finite group. Let p be an odd prime and n be even. Then $\pi_i(S^n)$ is \mathcal{C}_p -isomorphic to $\pi_{i-1}(S^{n-1}) + \pi_i(S^{2n-1})$. Let n be odd. Then $\pi_{n+k}(S^n) \in \mathcal{C}_p$ ($k < 2p - 3$), and $\pi_{n+2p-3}(S^n)$ is \mathcal{C}_p -isomorphic to \mathbb{Z}_p . Serre and H. Toda determined $\pi_{n+k}(S^n)$ for $k = 3, 4, 5$, and Serre further determined it for $k = 6, 7, 8$. Utilizing the reduced product space of S^n , James gave the sequence

$$\begin{aligned} \dots \rightarrow \pi_i(S^n) \xrightarrow{E} \pi_{i+1}(S^{n+1}) \xrightarrow{H} \pi_{i+1}(S^{2n+1}) \\ \rightarrow \pi_{i-1}(S^n) \rightarrow \dots \end{aligned}$$

and showed that it is an exact sequence if n is odd and an exact sequence mod 2 if n is even (1953). Using this exact sequence and the secondary composition, Toda determined $\pi_{n+k}(S^n)$ for $k \leq 19$ (\rightarrow Appendix A, Table 6.VI).

By the Freudenthal theorem (1), the $\pi_{n+k}(S^n)$ ($n > k + 1$) for a fixed k are isomorphic to each other. We call $\pi_{n+k}(S^n)$ ($n > k + 1$) the **stable homotopy group of the k -stem** of the sphere and denote it by G_k . For $k = 0, 1, 2, \dots, 15, \dots$, $G_k \cong \mathbb{Z}, \mathbb{Z}_2, \mathbb{Z}_2, \mathbb{Z}_{24}, 0, 0, \mathbb{Z}_2, \mathbb{Z}_{240}, \mathbb{Z}_2 + \mathbb{Z}_2, \mathbb{Z}_2 + \mathbb{Z}_2 + \mathbb{Z}_2, \mathbb{Z}_6, \mathbb{Z}_{504}, 0, \mathbb{Z}_3, \mathbb{Z}_2 + \mathbb{Z}_2, \mathbb{Z}_{480} + \mathbb{Z}_2, \dots$. For the computation of G_k , the notion of n -connective fiber spaces is important. By

utilizing the Adams spectral sequence, we can show that G_k is closely related to the cohomology of the \dagger Steenrod algebra. Let p be an odd prime. There exist the following sequences of elements of order p : $\{\alpha_i \in G_k (k = 2i(p-1) - 1)\}$, $\{\beta_i \in G_k (k = 2(ip + i - 1)(p-1) - 2)\}$ and for $p > 3$ $\{\gamma_i \in G_k (k = 2(ip^2 + (i-1)p + i - 2)(p-1) - 3)\}$. The p -component of G_k is determined for $k < 2p^2(p-1) - 3$ by using Steenrod algebra. To compute G_k for higher k , relations such as $\alpha_1 \beta_1^p = 0$, $\beta_i \beta_1^p = 0$ ($i > 1$) are necessary. In general, each element of G_k ($k \neq 0$) is nilpotent (G. Nishida). Let $\pi_i(S^n; p)$ be the p -component of $\pi_i(S^n)$. To survey this group for the nonstable case ($i \geq 2n - 1$), we utilize Serre's mod p direct sum decomposition (for n even), and we have the following two exact sequences for the case of odd n :

$$\begin{aligned} \dots \rightarrow \pi_i(S^n) \xrightarrow{E^2} \pi_{i+2}(S^{n+2}) \rightarrow \pi_i(\Omega^2(S^{n+2}), S^n) \xrightarrow{\circ} \dots, \\ \dots \rightarrow \pi_{i+3}(S^{pn+p+1}; p) \xrightarrow{\Delta} \pi_{i+1}(S^{pn+p-1}; p) \\ \rightarrow \pi_i(\Omega^2(S^{n+2}), S^n; p) \rightarrow \pi_{i+2}(S^{pn+p+1}; p) \xrightarrow{\Delta} \dots, \end{aligned}$$

where $E^2 = E \circ E$ and $\Delta E^2(\alpha) = p\alpha$ (\rightarrow Appendix A, Table 6.VI).

V. Homotopy Groups of Classical Groups

Consider the classical group $U(n, \Lambda)$, which is either the orthogonal group $O(n)$ ($\Lambda = \mathbf{R}$); the unitary group $U(n)$ ($\Lambda = \mathbf{C}$); or the symplectic group $Sp(n)$ ($\Lambda = \mathbf{H}$). The **infinite classical group** $U(\infty, \Lambda)$ is defined to be the inductive limit group of $\{U(n, \Lambda) | n = 1, 2, \dots\}$ with respect to the natural injection $U(n, \Lambda) \subset U(n+1, \Lambda)$. We call $U(\infty, \Lambda)$ the **infinite orthogonal group**, **infinite unitary group**, and **infinite symplectic group** for $\Lambda = \mathbf{R}$, \mathbf{C} , and \mathbf{H} , respectively. The dimensions of the cells of $U(\infty, \Lambda) - U(n, \Lambda)$ are $\geq \lambda(n+1) - 1$, where $\lambda = \dim_{\mathbf{R}} \Lambda$ ($= 1$ ($\Lambda = \mathbf{R}$), $= 2$ ($\Lambda = \mathbf{C}$), $= 4$ ($\Lambda = \mathbf{H}$)). It follows that $\pi_k(U(n, \Lambda))$ is isomorphic to $\pi_k(U(\infty, \Lambda))$ for $k < \lambda(n+1) - 2$, which is called the k th **stable homotopy group of the classical group**. Let $\mathbf{O} = U(\infty, \mathbf{R})$, $\mathbf{U} = U(\infty, \mathbf{C})$, $\mathbf{Sp} = U(\infty, \mathbf{H})$. The homotopy groups of the classical groups are periodic ($k \geq 0$):

$$\begin{aligned} \pi_k(\mathbf{U}) \cong \pi_{k+2}(\mathbf{U}) \cong \mathbf{Z}, \quad k \text{ odd}, \\ \cong 0, \quad k \text{ even}, \\ \pi_k(\mathbf{O}) \cong \pi_{k+4}(\mathbf{Sp}) \cong \pi_{k+8}(\mathbf{O}), \\ \cong \mathbf{Z}, \quad k \equiv 3, 7 \pmod{8}, \\ \cong \mathbf{Z}_2, \quad k \equiv 0, 1 \pmod{8}, \\ \cong 0, \quad k \equiv 0, 1, 3, 7 \pmod{8}. \end{aligned}$$

This is called the **Bott periodicity theorem**. The relations are deduced from weak homotopy equivalences $\mathbf{U} \rightarrow \Omega(B_{\mathbf{U}})$, $B_{\mathbf{U}} \times \mathbf{Z} \rightarrow \Omega(\mathbf{U})$, $B_{\mathbf{O}} \times \mathbf{Z} \rightarrow \Omega(\mathbf{U}/\mathbf{O})$, $\mathbf{U}/\mathbf{O} \rightarrow \Omega(\mathbf{Sp}/\mathbf{U})$, $\mathbf{Sp}/\mathbf{U} \rightarrow \Omega(\mathbf{Sp})$, \mathbf{Sp}

$\rightarrow \Omega(B_{\mathbf{Sp}})$, $B_{\mathbf{Sp}} \times \mathbf{Z} \rightarrow \Omega(\mathbf{U}/\mathbf{Sp})$, $\mathbf{U}/\mathbf{Sp} \rightarrow \Omega(\mathbf{O}/\mathbf{U})$, $\mathbf{O}/\mathbf{U} \rightarrow \Omega(\mathbf{O})$. This result is applied to nonstable cases; for example, $\pi_{2n}(U(n))$ is a cyclic group of order $n!$ (\rightarrow Appendix A, Table 6.VI). The 2-dimensional homotopy group $\pi_2(G)$ of any Lie group is trivial.

Let $\alpha \in \pi_k(O(n))$, where $\alpha = [f]$, $f: S^k \rightarrow O(n)$. We define $\bar{f}: S^k \times S^{n-1} \rightarrow S^{n-1}$ by $\bar{f}(x, y) = f(x) \cdot y$ and identify S^{k+n} with the boundary $(E^{k+1} \times S^{n-1}) \cup (S^k \times E^n)$ of $E^{k+1} \times E^n$. We extend \bar{f} to $f: S^{k+n} \rightarrow S^n$ so that it maps $E^{k+1} \times S^{n-1}$, $S^k \times E^n$ into the upper and lower hemisphere of S^n (S^{n-1} = the equator), respectively. Let $J(\alpha) \in \pi_{n+k}(S^n)$ be the class of the mapping thus obtained. This homomorphism $J: \pi_k(O(n)) \rightarrow \pi_{n+k}(S^n)$ is called a **J -homomorphism** of Hopf and Whitehead. For the stable case, $J: \pi_k(\mathbf{O}) \rightarrow G_k$ is injective for $k \equiv 0, 1 \pmod{8}$, and the order of the image of J is the denominator of $B_{2t}/4t$ (B_{2t} is a \dagger Bernoulli number) or its double for $k = 4t - 1$ (Adams).

References

- [1] H. Hopf, Über die Abbildungen von Sphären auf Sphären niedriger Dimension, Fund. Math., 25 (1935), 427–440.
- [2] W. Hurewicz, Beiträge zur Topologie der Deformationen I–IV, Proc. Acad. Amsterdam, 38 (1935), 112–229, 521–528; 39 (1936), 117–125, 215–224.
- [3] H. Freudenthal, Über die Klassen der Sphärenabbildungen, Compositio Math., 5 (1937), 299–314.
- [4] J.-P. Serre, Groupes d'homotopie et classes des groupes abéliens, Ann. Math., (2) 58 (1953), 258–294.
- [5] R. Bott, The stable homotopy of the classical groups, Ann. Math., (2) 70 (1959), 313–337.
- [6] H. Toda, Composition methods in homotopy groups of spheres, Princeton Univ. Press, 1962.
- [7] E. H. Spanier, Algebraic topology, McGraw-Hill, 1966.
- [8] R. M. Switzer, Algebraic topology, Springer, 1975.
- [9] G. W. Whitehead, Elements of homotopy theory, Springer, 1978.

203 (III.25) Hopf Algebras

A. General Remarks

The concept of Hopf algebras arose from two directions. First in the field of algebraic topol-

ogy the notion of Hopf algebras arose from the study of homology and cohomology of Lie groups or, more generally, H -spaces. It was introduced by H. Hopf [1], whose basic structure theorem was generalized and applied to several problems by A. Borel [2]. These Hopf algebras are graded as algebras and coalgebras, and they are now used as standard tools in algebraic topology. On the other hand, Hopf algebras without grading were studied in connection with affine algebraic groups and formal groups. The study of nongraded Hopf algebras as an algebraic system was initiated by M. E. Sweedler [3], and many results on Hopf algebras of this type have been applied, not only to the theory of algebraic groups but also to the Galois theory of field extension and to combinatorial theory.

Although these two types of Hopf algebras have similar structures, and the same terminology is used to describe their properties, they are somewhat different from each other. So to avoid confusion in this article, we distinguish between graded Hopf algebras and Hopf algebras.

B. Graded Algebras

A \mathbb{Z} -graded module $A = \sum_{n \geq 0} A_n$ over a field k is said to be of **finite type** when each A_n is finite-dimensional, A is **connected** when an isomorphism $\eta: k \cong A_0$ is given. The \mathbb{Z} -tensor product of two graded modules A and B is a graded module with $A \otimes B = \sum_n (A \otimes B)_n$, $(A \otimes B)_n = \sum_p A_p \otimes B_{n-p}$. We call $A^* = \sum A_n^*$ (where A_n^* is the \mathbb{Z} -dual module of A_n) the **dual graded module** of A . When A and B are of finite type, $A \otimes B$ and A^* are also of finite type, and we have $(A \otimes B)^* = A^* \otimes B^*$ and $A^{**} = A$. When A and B are connected, A^* and $A \otimes B$ are also connected.

Let A be a graded module. If there exists a degree-preserving linear mapping $\varphi: A \otimes A \rightarrow A$, we call (A, φ) a **graded algebra**, whereas if there exists a degree-preserving linear mapping $\psi: A \rightarrow A \otimes A$, we call (A, ψ) a **graded coalgebra**. We call φ a **multiplication**, and ψ a **comultiplication** (or **diagonal mapping**). Usually we write $\varphi(a \otimes b) = ab$ (the **product** of $a, b \in A$), and call $\psi(a)$ the **coproduct** of a . Multiplication and comultiplication are dual operations. If A is of finite type and (A, φ) is a graded algebra, then (A^*, φ^*) (where φ^* is the dual mapping of φ) is a graded coalgebra, and vice versa. A multiplication φ is called **associative** (**commutative**) if $\varphi(1 \otimes \varphi) = \varphi(\varphi \otimes 1)$ ($\varphi \circ T = \varphi$), where $T: A \otimes A \rightarrow A \otimes A$ is the mapping defined by $T(a \otimes b) = (-1)^{pq} b \otimes a$ for $a \in A_p$ and $b \in A_q$. Associativity and commutativity of a comultiplication are defined dually. Let (A, ψ) be a

graded connected coalgebra of finite type, and identify k with A_0 via η . Then the graded algebra (A^*, ψ^*) has the unity of k as unity if and only if ψ satisfies $\psi(1) = 1 \otimes 1$ and $\psi(x) = 1 \otimes x + x \otimes 1 + \sum_i x'_i \otimes x''_i$ ($0 < \deg x'_i < \deg x$) for $\deg x > 0$. In this case we say that ψ has the unity of k as **counity**. For graded algebras (A, φ) and (B, φ') , if $\varphi'' = (\varphi \otimes \varphi') \circ (1 \otimes T \otimes 1): A \otimes B \otimes A \otimes B \rightarrow A \otimes B$, then $(A \otimes B, \varphi'')$ is also a graded algebra, which we denote by $(A \otimes B, \varphi'') = (A, \varphi) \otimes (B, \varphi')$. The tensor product of graded coalgebras is defined as the dual notion of $(A, \varphi) \otimes (B, \varphi')$.

C. Graded Hopf Algebras

For simplicity we assume that graded modules are defined over a field, connected, and of finite type. Let a graded module A be equipped with a multiplication φ and a comultiplication ψ . If φ and ψ have the unity of k as unity and $\psi: (A, \varphi) \rightarrow (A, \varphi) \otimes (A, \varphi)$ is an algebra homomorphism, then we call (A, φ, ψ) a **graded Hopf algebra**. The last condition for a graded Hopf algebra is satisfied if and only if $\varphi: (A, \psi) \otimes (A, \psi) \rightarrow (A, \psi)$ is a homomorphism of graded coalgebras. The dual (A^*, ψ^*, φ^*) is also a graded Hopf algebra, called the **dual Hopf algebra** of (A, φ, ψ) .

D. H -Spaces

Let X be a topological space. The \mathbb{Z} -cohomology group $H^*(X)$ (\mathbb{Z} -homology group $H_*(X)$) considered over a field k has a multiplication d^* (comultiplication d_*), which is induced by the diagonal mapping $d: X \rightarrow X \times X$ and becomes a commutative and associative graded algebra (coalgebra). The groups $H^*(X)$ and $H_*(X)$ are dual to each other (\rightarrow 201 Homology Theory I, J). When X is equipped with a base point x_0 and a base point-preserving continuous mapping $h: X \times X \rightarrow X$ such that $h \circ \iota_i \simeq 1_X$ (\mathbb{Z} -homotopic) for $i = 1$ and 2 (where $\iota_1(x) = (x, x_0)$ and $\iota_2(x) = (x_0, x)$), we call (X, h) an **H -space**, h a **multiplication**, and x_0 a **homotopy identity** of X . Then h induces, through a \mathbb{Z} -K nneth isomorphism, a comultiplication $h^*: H^*(X) \rightarrow H^*(X) \otimes H^*(X)$ (**Hopf comultiplication**) and a multiplication $h_*: H_*(X) \otimes H_*(X) \rightarrow H_*(X)$ (**Pontryagin multiplication**). Then $h^*(\alpha)$ ($\alpha \in H^*(X)$) is called the **Hopf coproduct** of α , and $h_*(\beta \otimes \gamma)$ ($\beta, \gamma \in H_*(X)$) is called the **Pontryagin product** of β and γ . When X is \mathbb{Z} -arcwise connected and $H_*(X)$ is of finite type, $(H^*(X), d^*, h^*)$ and $(H_*(X), h_*, d_*)$ are graded Hopf algebras dual to each other. In particular, when h is **homotopy associative**, i.e., $h \circ (h \times 1_X) \simeq h \circ (1_X \times h)$ (**homotopy commutative**, i.e., $h \simeq h \circ T$, where

$T(x_1, x_2) = (x_2, x_1)$ for $x_i \in X$), then h^* and h_* are associative (commutative). $^* \text{Topological groups and } ^* \text{loop spaces are homotopy associative } H\text{-spaces. If a continuous mapping } g: X \rightarrow X \text{ satisfies } h \circ (1_X \times g) \cong h \circ (g \times 1_X) \cong c \text{ (constant mapping } X \rightarrow \{x_0\}), \text{ then } g \text{ is called a homotopy inverse for } X, h.$

Suppose that a graded Hopf algebra A is defined over a field k of characteristic p and equipped with associative and commutative multiplication, and A is generated by a single element $a \in A_n$. Then A is a $^* \text{polynomial ring } k[a]$ (n is even when $p \neq 2$) or a $^* \text{quotient ring } k[a]/(a^2)$ (n is odd when $p \neq 2$) or $k[a]/(a^{p^f})$ (only when $p \neq 0$; n is even when $p \neq 2$). These are called **elementary Hopf algebras**. Every graded Hopf algebra over a $^* \text{perfect field } k \text{ with associative and commutative multiplication is isomorphic (as a graded algebra) to a tensor product of elementary Hopf algebras (Borel's theorem) [2]. In particular, the cohomology algebra over a field of characteristic 0 of a } ^* \text{compact connected Lie group is isomorphic to a } ^* \text{Grassmann algebra generated by elements of odd degrees [1].$

E. Steenrod Algebras

The $^* \text{Steenrod algebra } \mathcal{A}_p \text{ over } \mathbf{Z}_p \text{ is generated by } ^* \text{Steenrod operations } Sq^i \text{ } (p=2), \mathcal{P}^i \text{ } (p>2), \text{ and the } ^* \text{Bockstein operation } \Delta_p \text{ } (p>2), \text{ with composition of operations defined as multiplication. Then } \mathcal{A}_p \text{ is a connected associative graded algebra of finite type (not commutative). Defining a comultiplication } \psi \text{ of } \mathcal{A}_p \text{ by } \psi(Sq^n) = \sum Sq^i \otimes Sq^{n-i}, \psi(\mathcal{P}^n) = \sum \mathcal{P}^i \otimes \mathcal{P}^{n-i}, \text{ and } \psi(\Delta_p) = 1 \otimes \Delta_p + \Delta_p \otimes 1, \mathcal{A}_p \text{ becomes a graded Hopf algebra with an associative and commutative comultiplication. Thus its dual } \mathcal{A}_p^* \text{ is a graded Hopf algebra with an associative and commutative multiplication, and we can apply Borel's theorem to } \mathcal{A}_p^* \text{ in order to investigate the structure of } \mathcal{A}_p \text{ [4].}$

Let (A, φ, ψ) be a graded Hopf algebra with associative multiplication and comultiplication. Putting $c(1) = 1$ and $c(a) = -a - \sum a'_i \cdot c(a''_i)$ for $\deg a > 0$ (where $\psi(a) = 1 \otimes a + a \otimes 1 + \sum a'_i \otimes a''_i$), we obtain a linear mapping $c: A \rightarrow A$ satisfying $c\varphi = \varphi(c \otimes c)T$. We call c the **conjugation mapping** of A . When the multiplication or comultiplication is commutative, we obtain the relation $c^2 = 1$, and c is a bijection. The conjugation mapping is utilized in studying Steenrod algebras [4, 5].

F. Coalgebras

Now we turn to nongraded cases. Let A be a vector space over a field k , and let $\mu: A \otimes_k A \rightarrow A$ and $\eta: k \rightarrow A$ be linear mappings. Then the

triple (A, μ, η) is said to be an **algebra** over k if $\mu \circ (\mu \otimes 1_A) = \mu \circ (1_A \otimes \mu)$ and $\mu \circ (1_A \otimes \eta) = \mu \circ (\eta \otimes 1_A) = 1_A$, where 1_A is the identity mapping of A , and $A \otimes_k k$ and $k \otimes_k A$ are identified naturally with A . We call μ the **multiplication** and η the **unit mapping** of the algebra. Dually a triple (C, Δ, ε) with C a vector space over k , linear mappings $\Delta: C \rightarrow C \otimes_k C$, and $\varepsilon: C \rightarrow k$ is said to be a **coalgebra** over k if $(1_C \otimes \Delta) \circ \Delta = (\Delta \otimes 1_C) \circ \Delta$ and $(1_C \otimes \varepsilon) \circ \Delta = (\varepsilon \otimes 1_C) \circ \Delta = 1_C$. We call Δ the **comultiplication** or the **diagonal mapping** and ε the **augmentation** or the **counit** of the coalgebra.

An algebra (A, μ, η) is called **commutative** if $\mu \circ T = \mu$, where T is the twist mapping $a \otimes b \mapsto b \otimes a$. A **cocommutative coalgebra** is defined dually. Definitions of the tensor product of two algebras or coalgebras are similar to the definitions in the graded case. If D is a subspace of a coalgebra (C, Δ, ε) over k satisfying $\Delta(D) \subset D \otimes_k D$, then $(D, \Delta|_D, \varepsilon|_D)$ is a coalgebra and is said to be a **subcoalgebra** of C . A subspace I of a coalgebra (C, Δ, ε) over k is called a **coideal** of C if $\Delta(I) \subset C \otimes_k I + I \otimes_k C$ and $\varepsilon(I) = 0$. Then the quotient space C/I has a coalgebra structure induced naturally from (C, Δ, ε) and is said to be a **quotient coalgebra** of C . The intersection and the sum of subcoalgebras of C are again subcoalgebras of C . If S is a subset of C , then the intersection of all subcoalgebras containing S is said to be the subcoalgebra generated by S . The subcoalgebra generated by any finite set or finite-dimensional subspace of C is finite-dimensional.

If (C, Δ, ε) is a coalgebra over k , then the dual space C^* of C has an algebra structure over k with multiplication μ and unit mapping η defined naturally from the $^* \text{dual mappings of } \Delta \text{ and } \varepsilon \text{ respectively. } (C^*, \mu, \eta) \text{ is called the dual algebra of } (C, \Delta, \varepsilon). \text{ Suppose that } (A, \mu, \eta) \text{ is an algebra over } k, \text{ and let } A^\circ \text{ be the subset of the dual space } A^* \text{ of } A \text{ consisting of elements } f \text{ whose kernel contains an ideal } I \text{ such that } A/I \text{ is finite-dimensional. Then } (A^\circ, \Delta, \varepsilon) \text{ is a coalgebra over } k, \text{ where } \Delta \text{ and } \varepsilon \text{ are the linear mappings induced from the dual ones of } \mu \text{ and } \eta, \text{ respectively. } (A^\circ, \Delta, \varepsilon) \text{ is called the dual coalgebra of } (A, \mu, \eta). \text{ The } ^* \text{functors } ()^* \text{ and } ()^\circ \text{ are adjoint to one another in the sense that there is a natural bijective correspondence between the set of algebra homomorphisms of } A \text{ to } C^* \text{ and that of coalgebra homomorphisms of } C \text{ to } A^\circ \text{ for any coalgebra } C \text{ and algebra } A, \text{ where coalgebra homomorphisms are defined as the dual notion of } ^* \text{algebra homomorphisms.}$

A nonzero subcoalgebra D of a coalgebra C is called **simple** if D has no nonzero proper subcoalgebras, and the sum of all simple subcoalgebras of C is called the **coradical** of C . If

C coincides with its coradical, then C is said to be **cosemisimple**. If C has only one simple subcoalgebra, then C is called **irreducible**. C is called **pointed** if all simple subcoalgebras of C are 1-dimensional. An element g of a coalgebra (C, Δ, ε) is called **grouplike** if $\Delta(g) = g \otimes g$ and $\varepsilon(g) = 1$. The set $G(C)$ of grouplike elements in C is linearly independent over k .

G. Bialgebras

A system $(H, \mu, \eta, \Delta, \varepsilon)$ with an algebra structure (H, μ, η) and a coalgebra structure (H, Δ, ε) is said to be a **bialgebra** over k if Δ and ε are algebra homomorphisms. This last condition is equivalent to saying that μ and η are coalgebra homomorphisms. If K is a subspace of a bialgebra $(H, \mu, \eta, \Delta, \varepsilon)$ which is simultaneously a subalgebra and a subcoalgebra of H , then we call K a **subbialgebra** of H . An ideal I of (H, μ, η) which is also a coideal of (H, Δ, ε) is called a **biideal** of H and the quotient space H/I has a bialgebra structure which is said to be a **quotient bialgebra** of H . A linear mapping between bialgebras is a **bialgebra homomorphism** if it is simultaneously an algebra homomorphism and a coalgebra homomorphism. A bialgebra $(H, \mu, \eta, \Delta, \varepsilon)$ is called commutative (cocommutative) if (H, μ, η) ((H, Δ, ε)) is commutative (cocommutative).

Examples. Let kG be the vector space with a set G as free basis over a field k . If we define linear mappings $\Delta: kG \rightarrow G \otimes_k kG$ by $\Delta(x) = x \otimes x$ and $\varepsilon: kG \rightarrow k$ by $\varepsilon(x) = 1$ for x in G , then $(kG, \Delta, \varepsilon)$ is a cocommutative coalgebra over k such that the set $G(kG)$ of grouplike elements is equal to G . Moreover if G is a \dagger semigroup with unit element, then $(kG, \mu, \eta, \Delta, \varepsilon)$ is a cocommutative bialgebra over k , where $\mu(\eta)$ is the multiplication (unit mapping) of the \dagger semigroup algebra kG . This bialgebra is called a **semigroup bialgebra** over k . If L is a \dagger Lie algebra over k and $U(L)$ the \dagger universal enveloping algebra of L with multiplication μ and unit mapping η , then Lie algebra homomorphisms $x \rightarrow x \oplus x$ ($L \rightarrow L \oplus L$) and $x \rightarrow 0$ ($L \rightarrow \{0\}$) induce algebra homomorphisms $\Delta: U(L) \rightarrow U(L \oplus L) \cong U(L) \otimes U(L)$ and $\varepsilon: U(L) \rightarrow k$, respectively. Then $(U(L), \mu, \eta, \Delta, \varepsilon)$ is a cocommutative bialgebra over k , called the **universal enveloping bialgebra** of L .

H. Hopf Algebras

Let (A, μ, η) be an algebra over k and (C, Δ, ε) a coalgebra over k . If f and g are in $R = \text{Hom}_k(C, A)$, then $f * g = \mu \circ (f \otimes g) \circ \Delta$ is called the **convolution** of f and g . Defining $\mu': R \otimes_k R \rightarrow R$ and $\eta': k \rightarrow R$ by $\mu'(f \otimes g) = f * g$

and $\eta'(\alpha) = \alpha \eta \circ \varepsilon$, then (R, μ', η') is an algebra over k . If H is a bialgebra over k with underlying coalgebra H^c and algebra H^A , then $R_H = \text{Hom}_k(H^c, H^A)$ is an algebra over k in the same manner as above. If the identity mapping 1_H of H has the inverse S in R_H under the multiplication μ' , then H is said to be a **Hopf algebra** over k with **antipode** S . Then S satisfies the following: $S(gh) = S(h)S(g)$ for g, h in H , $S \circ \eta = \eta$, $\varepsilon \circ S = \varepsilon$ and $T \circ (S \otimes S) \circ \Delta = \Delta \circ S$, where T is the twist mapping $a \otimes b \rightarrow b \otimes a$. If H is commutative or cocommutative, then $S \circ S = 1_H$.

If H' is another Hopf algebra over k with antipode S' , then a bialgebra homomorphism f of H to H' such that $S'f = fS$ is called a **Hopf algebra homomorphism**. If H and H' are both commutative (cocommutative), then any bialgebra homomorphism of H to H' is a Hopf algebra homomorphism. The \dagger category whose objects are commutative and cocommutative Hopf algebras over a field k and whose morphisms are Hopf algebra homomorphisms is an \dagger Abelian category (A. Grothendieck). If H is a commutative Hopf algebra over a field of characteristic zero, then the underlying algebra H^A has no \dagger nilpotent elements (P. Cartier).

If G is a \dagger group, then the group bialgebra kG given above has an antipode S defined by $S(x) = x^{-1}$ for x in G , and hence kG is a cocommutative Hopf algebra over k . Another example of Hopf algebras is the \dagger coordinate ring of an \dagger algebraic group defined over k . More generally, let $X = \text{Spec}(A)$ be an \dagger affine group scheme over k . Then algebra homomorphisms $\Delta: A \rightarrow A \otimes_k A$, $\varepsilon: A \rightarrow k$, and $S: A \rightarrow A$ are naturally induced from the group structure of X , and $(A, \mu, \eta, \Delta, \varepsilon, S)$ is a commutative Hopf algebra over k , where μ and η are the multiplication and unit mapping of A , respectively. Conversely, if $(A, \mu, \eta, \Delta, \varepsilon, S)$ is a commutative Hopf algebra over k , then $X = \text{Spec}(A)$ is an affine group scheme over k with the group structure induced from Δ, ε , and S . Hence a commutative Hopf algebra over k is nothing but a cogroup object of the category of commutative algebras over k (i.e., a \dagger group object of the \dagger dual category). Dually a cocommutative Hopf algebra over k is nothing but a group object of the category of cocommutative coalgebras over k .

I. Hyperalgebras

If (C, Δ, ε) is a pointed irreducible coalgebra over k , then C contains a unique grouplike element g , and kg is the unique simple subcoalgebra of C . An element a of C satisfying $\Delta(a) = a \otimes g + g \otimes a$ is called **primitive**. The set $P(C)$ of primitive elements in C is a vector

subspace of C . A cocommutative coalgebra C is called **colocal** if C is irreducible, i.e., if the dual algebra C^* of C is a \dagger quasilocal ring. The \dagger dimension of $P(C)$ of a colocal pointed coalgebra C is finite if and only if the dual algebra C^* of C is a \dagger Noetherian complete local ring. A bialgebra $(H, \mu, \eta, \Delta, \epsilon)$ over k is said to be a **hyperalgebra** over k if the underlying coalgebra is colocal. Then the unique simple subcoalgebra (grouplike element) of H is $\eta(k)(\eta(1))$, and $P(H)$ has a Lie algebra structure defined by $[x, y] = xy - yx$ for x, y in $P(H)$.

The universal enveloping bialgebra $U(L)$ of a Lie algebra L over k is a hyperalgebra over k such that the set $P(U(L))$ of primitive elements in $U(L)$ is equal to L . Conversely, if the characteristic of k is zero, any hyperalgebra H over k is isomorphic to the universal enveloping algebra $U(P(H))$ of the Lie algebra $P(H)$. But in positive-characteristic cases $U(P(H))$ is generally a proper subbialgebra of H . Another important example of hyperalgebras is the dual coalgebra $hy(X) = A^\circ$ of the \dagger stalk A at the neutral point of the \dagger structure sheaf of an \dagger algebraic group scheme X over k . In addition, $hy(X)$ has an algebra structure defined from the group structure of X and is a hyperalgebra over k such that $P(hy(X))$ is equal to the Lie algebra $L(X)$ of X . Although $L(X)$ plays an important role in the infinitesimal theory of algebraic groups over a field of characteristic zero, it does not give any information on infinitesimals of orders higher than p in the case of positive characteristic p . In the case of characteristic zero we see $hy(X) = U(L(X))$ and $L(X) = P(hy(X))$, and so $hy(X)$ is a natural substitute for $L(X)$ in positive-characteristic cases. From this viewpoint many interesting results on $hy(X)$ of an algebraic group scheme X which are parallel to those on $L(X)$ in the case of characteristic zero have been obtained [6–8].

References

- [1] H. Hopf, Über die Topologie der Gruppen-Mannigfaltigkeiten und ihre Verallgemeinerungen, Ann. Math., (2) 42 (1941), 22–52.
- [2] A. Borel, Sur la cohomologie des espaces fibrés principaux et des espaces homogènes des groupes de Lie compacts, Ann. Math., (2) 57 (1953), 115–207.
- [3] M. E. Sweedler, Hopf algebras, Benjamin, 1969.
- [4] J. W. Milnor, The Steenrod algebra and its dual, Ann. Math., (2) 67 (1958), 150–171.
- [5] J. W. Milnor and J. C. Moore, On the structure of Hopf algebras, Ann. Math., (2) 81 (1965), 211–264.

- [6] J. Dieudonné, Introduction to the theory of formal groups, Dekker, 1973.
- [7] M. Takeuchi, Tangent coalgebras and hyperalgebras I, Japan. J. Math., 42 (1974), 1–143.
- [8] H. Yanagihara, Theory of Hopf algebras attached to group schemes, Lecture notes in math. 614, Springer, 1977.
- [9] E. Abe, Hopf algebras, Cambridge Univ. Press, 1980. (Original in Japanese, 1977.)
- [10] G.-C. Rota, Coalgebras and bialgebras in combinatorics, Lecture notes at the umbral calculus conference, University of Oklahoma, 1978.

204 (XX.10) Hydrodynamical Equations

A. General Remarks

Mathematical analysis of the motion of fluids (\rightarrow 205 Hydrodynamics) gives rise to various kinds of mathematical problems; the equations that govern flows are amongst the most important and most extensively studied of the nonlinear partial differential equations. Here we review only basic and noteworthy results. Sections B–E are concerned with incompressible fluids, while Sections F and G deal with compressible fluids.

B. Nonstationary Solutions of the Navier-Stokes Equation

Let Ω be a bounded domain in \mathbf{R}^m ($m = 2$ or 3) occupied by a fluid, with smooth boundary $\partial\Omega$. If the fluid is viscous and incompressible, its motion can be described by means of the velocity $u = u(t, x)$ and the pressure $p = p(t, x)$ with $t \geq 0$ and $x \in \Omega$ the time variable and the space variable, respectively. For simplicity, we assume that external forces are absent. Then u and p satisfy the **Navier-Stokes equation**

$$\frac{\partial u}{\partial t} + (u \nabla)u = \nu \Delta u - \nabla p, \quad (1)$$

and the **equation of continuity**

$$\operatorname{div} u = 0, \quad (2)$$

where the positive constant ν stands for the (kinetic) viscosity. On $\partial\Omega$, u is subject to the boundary condition

$$u|_{\partial\Omega} = \beta(t, x). \quad (3)$$

The initial value of u is also prescribed:

$$u|_{t=0} = u_0(x). \quad (4)$$

If Ω is unbounded, e.g., if Ω is an exterior domain outside compact surfaces, then

$$u(t, x) \rightarrow U_0(t) \quad (|x| \rightarrow \infty) \quad (3')$$

is imposed in addition as the boundary condition at infinity.

The problem of finding u and p that satisfy (1)–(4) for given β and u_0 is called the **Navier-Stokes initial value problem** (abbreviation, NS initial value problem). Pioneering mathematical studies of this problem were initiated by J. Leray, and since the 1950s various contributions have been made by many authors, including E. Hopf, O. A. Ladyzhenskaya, H. Fujita, T. Kato, and S. Ito (\rightarrow [13, 17, 21]). We state here the result in terms of regular (classical) solutions under the simplifying assumption that Ω is bounded, $\beta \equiv 0$, and u_0 is solenoidal ($\operatorname{div} u_0 = 0$) and smooth. The situation depends nontrivially upon the dimension m . Namely, if $m = 2$, then a regular solution of the NS initial value problem exists uniquely and globally, i.e., for all time. If $m = 3$, we can prove only a local existence theorem (i.e., one holding in a finite interval) of a regular solution. This solution is unique in the interval of its existence; however, it can be extended over the whole interval when the Reynolds number is sufficiently small. In other words, the question of well-posedness of the 3-dimensional NS initial value problem is open at present.

C. Weak and Strong Solutions of the Navier-Stokes Equation

In 1951 Hopf introduced the notion of **weak solutions** of the NS initial value problem and succeeded in proving their global existence (without uniqueness). We here give the definition of **Hopf's weak solution**: Let $C_{0,\sigma}^\infty$ be the set of vector functions $u \in C_{0,\sigma}^\infty(\Omega)$ with $\operatorname{div} u = 0$. By H we denote the closure of $C_{0,\sigma}^\infty$ under the L^2 -norm, and by V the closure of $C_{0,\sigma}^\infty$ under the $W_2^1(\Omega)$ -norm (or equivalently, the Dirichlet norm for bounded Ω). Then an H -valued function $u = u(t)$ is a weak solution of the NS initial value problem with $\beta \equiv 0$ in $[0, T)$ ($0 < T \leq +\infty$) if

(i) $u \in L^\infty(0, T; H) \cap L^2(0, \infty; V)$,

(ii) u is weakly continuous from $[0, T)$ to H , and

(iii) u satisfies the weak equation

$$\int_0^T \left\{ (u, \varphi_t)_{L^2(\Omega)} - \nu (\nabla u, \nabla \varphi)_{L^2(\Omega)} + ((u \cdot \nabla) \varphi, u)_{L^2(\Omega)} \right\} dt = -(u_0, \varphi(0))_{L^2(\Omega)} \quad (5)$$

for all $\varphi \in C_0^\infty([0, T) \times \Omega)$ with $\operatorname{div} \varphi = 0$.

Note that the pressure p has been eliminated

in the weak equation. In general, the uniqueness of Hopf's weak solution is not known except for $m = 2$. However, when a regular solution does exist, then the weak solution coincides with it (a.e.), and is unique. Solutions of the NS initial value problem which are stronger than the weak solution to the extent that the uniqueness can be proved have been introduced, for instance, by A. A. Kiselev and Ladyzhenskaya [12] and Fujita and Kato [5]. Actually the existence and uniqueness theorems of the aforementioned regular solutions are proved using existence theorems of such **strong solutions**. Recently, Y. Giga and T. Miyakawa succeeded in generalizing the Fujita-Kato theory from L^2 to L^p , and thus they have shown that if $u_0 \in L^m(\Omega)$ and is solenoidal, a unique strong solution exists at least locally for any m .

Those strong solutions that satisfy uniqueness, and hence are regular solutions of the NS initial value problem, turn out to be smooth for $t > 0$; they are analytic in $t > 0$ and $x \in \Omega$.

D. Stationary Solutions of the Navier-Stokes Equation

If the flow is steady, u and p are solutions of the boundary value problem consisting of (1) with $\partial u / \partial t$ omitted, (2) and (3) with β independent of t (and, in addition, (3') with a constant U_0 if Ω is an exterior domain). The existence of solutions of this boundary value problem for the case of bounded Ω was established by Leray as one of the earliest applications of his fixed-point theorem. For the case of unbounded Ω , Leray's study was completed and extended by R. Finn, H. Fujita, and others to yield theorems on existence, regularity, and asymptotic behavior in the wakes of solutions (\rightarrow [4, 13, 21]). These stationary solutions are unique if the Reynolds number is sufficiently small. On the other hand, under certain circumstances that involve Quette flow, non-uniqueness or bifurcation of stationary solutions for large Reynolds numbers has been positively proved.

E. Euler's Equation

If the fluid is inviscid and incompressible, the Navier-Stokes equation is reduced to Euler's equation

$$\frac{\partial u}{\partial t} + (u \cdot \nabla) u = -\nabla p. \quad (6)$$

Then the boundary condition is replaced by the frictionless boundary condition, which, in

the homogeneous case, takes the form

$$u_n|_{\partial\Omega}=0, \quad (7)$$

where u_n is the normal component of u . Regular solutions of the initial value problem consisting of (6) with (2), (4), and (7) have been proved to exist for all t if $m=2$ and in a finite time if $m=3$ [1, 11].

F. The General Navier-Stokes Equations

If the fluid is compressible, viscous, and heat-conductive, its motion is described in terms of the density ρ , the velocity u , and some thermodynamic quantity, say the absolute temperature θ , and is governed by the following system of equations, sometimes called the **general Navier-Stokes equations**:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) = 0, \quad (8)$$

$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u = \frac{\mu}{\rho} \left(\Delta + \frac{1}{3} \nabla \operatorname{div} \right) u - \frac{1}{\rho} \nabla p, \quad (9)$$

$$\frac{\partial \theta}{\partial t} + (u \cdot \nabla)\theta = \frac{\kappa}{c_v \rho} \Delta \theta + \frac{\theta}{c_v \rho} (\operatorname{div} u) p_\theta + \frac{\Psi}{c_v \rho}. \quad (10)$$

Here the pressure p is regarded as a function of ρ and θ through the equation of state. The viscosity coefficient μ , the coefficient of heat conduction κ , and c_v , the specific heat at constant volume, are positive constants. We have for simplicity assumed that the external force is absent, and that the Stokes condition for viscosity is satisfied. Finally, Ψ is the dissipation function:

$$\Psi = -\frac{2}{3} \mu (\operatorname{div} u)^2 + \frac{\mu}{2} \sum_{i,j=1}^m \left(\frac{\partial}{\partial x_i} u_j + \frac{\partial}{\partial x_j} u_i \right)^2.$$

If Ω is the whole space, then the initial values ρ_0, u_0, θ_0 of ρ, u, θ are given at $t=0$, and we obtain the Cauchy problem for the general Navier-Stokes equation. Mathematical study of this Cauchy problem has become active since J. Nash [19] and N. Itaya [8] proved the existence of unique regular solutions local in time. Following Itaya's argument, A. Tani constructed a unique regular solution local in time for the initial boundary value problem consisting of (8)–(10) and boundary conditions imposed on u and θ . When it comes to the global existence of solutions, only restricted results are known. The global existence of a regular solution in the 1-dimensional version of the Cauchy problem has been established by Ya. Kanel' [10] and Itaya under certain simplifying assumptions, such that the fluid is a barotropic gas, i.e., one obeying $p = C\rho^\gamma$, where C and $\gamma \geq 1$ are positive constants. The

1-dimensional initial boundary value problem can also be solved globally if the gas is ideal and polytropic, i.e., one for which $p = R\rho\theta$, and with the internal energy proportional to θ and [9]. For the 3-dimensional case, we can only refer to [18], where existence of global solutions of the Cauchy problem has been proved for initial data close to constants under the assumption that the gas is ideal and polytropic.

G. Equations for Inviscid Ideal Gases

When the gas under consideration is inviscid, ideal, and barotropic, we put $\mu=0$ and $p = C\rho^\gamma$ in (9). The equation thus obtained is combined with (8) to yield the following quasilinear hyperbolic system, which admits conservation laws:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) &= 0, \\ \frac{\partial u}{\partial t} + (u \cdot \nabla)u + \frac{C}{\rho} \nabla \rho^\gamma &= 0. \end{aligned} \quad (11)$$

If the initial data ρ_0 and u_0 are smooth to some extent, then the Cauchy problem for (11) has a regular solution local in time. Generally, we cannot expect existence of regular solutions global in time, namely, discontinuity is likely to take place in a finite time, which corresponds to the occurrence of shock waves. Therefore we have to introduce weak solutions that admit discontinuity. By definition a piecewise continuous function $\{\rho, u\}$ is a weak solution of (11) if it satisfies (11) in the distribution sense and if its discontinuity is subject to a certain jump condition, called the **Rankine-Hugoniot relation** as well as another condition, called the **entropy condition**, which two conditions allow us to distinguish a physically realizable solution among many possible discontinuous solutions [3, 14]. Global existence of the weak solution has been proved so far only for the 1-dimensional problem with initial data close to constants in the sense that their oscillations and total variations are sufficiently small. Actually in this case we can apply J. Glimm's method [6] to construct weak solutions by means of a difference approximation which involves random numbers. Little is known regarding the uniqueness of weak solutions. Finally, if we are concerned with steady-state solutions of an inviscid compressible fluid and assume that the flow is irrotational, then we are led to a quasilinear partial differential equation of mixed type for the velocity potential Φ , which is elliptic in the subsonic region and hyperbolic in the supersonic region. Classical results concerning these equations may be found in [2, 3].

References

- [1] C. Bardos, Euler equation and Berger equation—Relation to turbulence, Lecture notes in math. 648, Springer, 1978, 1–46.
- [2] L. Bers, Mathematical aspects of subsonic and transonic gas dynamics, Wiley, 1958.
- [3] R. Courant and K. O. Friedrichs, Supersonic flow and shock waves, Interscience, 1948.
- [4] R. Finn, Stationary solutions of the Navier-Stokes equation, Amer. Math. Soc. Proc. Symp. Appl. Math., 17 (1965), 121–153.
- [5] H. Fujita and T. Kato, On the Navier-Stokes initial value problem I, Arch. Rational Mech. Anal., 16 (1964), 269–315.
- [6] J. Glimm, Solutions in the large for nonlinear hyperbolic systems of equations, Comm. Pure Appl. Math., 18 (1965), 697–715.
- [7] E. Hopf, Über die Anfangswertaufgabe für die hydrodynamischen Grundgleichungen, Math. Nachr., 4 (1950–1951), 213–231.
- [8] N. Itaya, On the Cauchy problem for the system of fundamental equations describing the movement of compressible viscous fluid, Kōdai Math. Sem. Rep., 23 (1971), 60–120; corrigenda and addenda, J. Math. Kyoto Univ., 16 (1976), 239–240.
- [9] A. V. Kazhikhov and V. V. Shelukhin, Overall solvability of single-valued functions for the initial boundary problem for one-dimensional equations of viscous gases (in Russian), Prikl. Mat. Mekh., 41 (1977), 282–291.
- [10] Ya. I. Kanel', A model system of equations for the movement of gases in one dimension (in Russian), Differentsial'nye Uravneniya, 4 (1968), 721–734.
- [11] T. Kato, On classical solutions of two-dimensional nonstationary Euler equation, Arch. Rational Mech. Anal., 25 (1967), 188–200.
- [12] A. A. Kiselev and O. A. Ladyzhenskaya, On the existence and uniqueness of solutions of nonstationary problems for viscous, incompressible fluids (in Russian), Izv. Akad. Nauk SSSR, 21 (1957), 655–680.
- [13] O. A. Ladyzhenskaya, The mathematical theory of viscous incompressible flows, Gordon & Breach, 1963. (Original in Russian, 1961.)
- [14] P. D. Lax, Hyperbolic systems of conservation laws and the mathematical theory of shock waves, SIAM Regional Conf. Ser. Appl. Math., 11 (1973).
- [15] J. Leray, Etude de diverses équations intégrales non linéaires et de quelques problèmes que pose l'hydrodynamique, J. Math. Pures Appl., (9) 12 (1933), 1–82.
- [16] J. Leray, Sur le mouvement d'un liquid visqueux emplissant l'espace, Acta Math., 63 (1934), 193–248.
- [17] J. L. Lions, Quelques methodes de résolution des problemes aux limites non linéaires, Dunod, 1969.
- [18] A. Matsumura and T. Nishida, The initial value problem for the equations of motion of viscous and heat-conductive gases, J. Math. Kyoto Univ., 20 (1980), 67–104.
- [19] J. Nash, Le problème de Cauchy pour les équations différentielles d'un fluide général, Bull. Soc. Math. France, 90 (1962), 487–497.
- [20] B. L. Rozhdestvenskii and N. N. Yanenko, Systems of quasilinear equations and their application to gas dynamics (in Russian), Nauka, 1978.
- [21] R. Temam, Navier-Stokes equations, North-Holland, 1977.

205 (XX.9) Hydrodynamics

A. General Remarks

Gases and liquids are easily deformed, and they share many kinetic properties. They are examples of fluids. By definition, a **fluid** is a continuous substance having the property that when it is not moving, any part of the substance separated from the rest by a surface exerts an outward force that is perpendicular to the given surface.

Hydrodynamics (or **fluid dynamics**) is concerned with the equilibrium and the motion of gases and liquids without considering their molecular structure. In particular, the branch of the theory concerning fluids in equilibrium is called **hydrostatics**, and **hydrodynamics** sometimes refers to the branch concerning fluids in motion.

There are two methods of describing the motion of a fluid. One regards a fluid as a system consisting of an infinite number of particles and discusses the motion of each particle as a function of time. This is **Lagrange's method**. For example, suppose that a fluid particle with the coordinates $(x, y, z) = (a, b, c)$ at the moment $t = 0$ has coordinates $x = f_1(a, b, c, t)$, $y = f_2(a, b, c, t)$, $z = f_3(a, b, c, t)$ at an arbitrary time t . Then the motion of the fluid is perfectly determined by the functions f_1 , f_2 , and f_3 .

The other is **Euler's method**, which discusses the values of the velocity $\mathbf{v}(u, v, w)$, the density ρ , the pressure p , etc., of the fluid at arbitrary times and positions. From this standpoint each quantity of the fluid is regarded as a function of a space-time point (x, y, z, t) .

The rate at which any physical quantity F varies while moving with the fluid particle is

the **Lagrangian derivative** DF/Dt , which is related to the ordinary partial derivatives by

$$\frac{DF}{Dt} = \frac{\partial F}{\partial t} + u \frac{\partial F}{\partial x} + v \frac{\partial F}{\partial y} + w \frac{\partial F}{\partial z}.$$

The three components (u, v, w) and the two state quantities (p, ρ) (in general, other state quantities, for example, the temperature T and the entropy S , are assumed to be determined by equations of state such as $T = T(p, \rho)$, $S = S(p, \rho)$) are determined by five ($= 1 + 3 + 1$) relations derived from the conservation laws of mass, momentum, and energy, namely, the **equation of continuity**, which corresponds to the conservation of mass,

$$\partial \rho / \partial t + \operatorname{div}(\rho \mathbf{v}) = 0; \quad (1)$$

the **equation of motion**, which corresponds to the conservation of momentum,

$$\partial(\rho \mathbf{v}) / \partial t + \operatorname{div}(\rho \mathbf{v} \otimes \mathbf{v} - \mathbf{p}) = \rho \mathbf{K}, \quad (2)$$

where \mathbf{K} is the external force per unit mass, \mathbf{p} is the stress tensor, and \otimes denotes the tensor product, while divergence is applied to each row vector, and by virtue of (1), the equation (2) can be expressed component-wise as

$$\begin{aligned} \rho \frac{Du}{Dt} &= \frac{\partial p_{xx}}{\partial x} + \frac{\partial p_{xy}}{\partial y} + \frac{\partial p_{xz}}{\partial z} + \rho K_x, \\ \rho \frac{Dv}{Dt} &= \frac{\partial p_{yx}}{\partial x} + \frac{\partial p_{yy}}{\partial y} + \frac{\partial p_{yz}}{\partial z} + \rho K_y, \\ \rho \frac{Dw}{Dt} &= \frac{\partial p_{zx}}{\partial x} + \frac{\partial p_{zy}}{\partial y} + \frac{\partial p_{zz}}{\partial z} + \rho K_z, \end{aligned} \quad (2')$$

and the **energy equation**, which corresponds to the conservation of energy,

$$\partial(\rho v^2/2 + \rho E) / \partial t + \operatorname{div}(\rho \mathbf{v}(v^2/2 + E) - \mathbf{v} \cdot \mathbf{p} + \mathbf{h}) = \rho \mathbf{v} \cdot \mathbf{K}, \quad (3)$$

or the **equation of entropy production**, which is another expression of (3),

$$\rho T DS/Dt = -\operatorname{div} \mathbf{h} + Q, \quad (3')$$

where E is the internal energy per unit mass, Q the heat generated per unit time and volume, and \mathbf{h} the heat flux. Here \mathbf{K} , p_{ik} , \mathbf{h} , and Q or their relations with other quantities (e.g., $\mathbf{h} = -\kappa \operatorname{grad} T$, where κ is the thermal conductivity) are assumed to be known.

B. Perfect Fluids

When there is a velocity gradient in the flow, a tangential stress appears which tends to make the velocity uniform, so that \mathbf{p} is not a diagonal tensor ($-p \delta_{ik}$, i.e., pressure). This property is called fluid **viscosity**. Generally, Q and \mathbf{h} do not vanish in this case. However, in order to simplify the problem we consider a

nonviscous (sometimes also adiabatic) fluid, which is called a **perfect fluid** and is a good approximation in the large of actual fluids. The motion of a perfect fluid is determined by **Euler's equation of motion**

$$\rho D\mathbf{v}/Dt = -\operatorname{grad} p + \rho \mathbf{K}, \quad (4)$$

which is obtained from (1) and (2) by replacing p_{ik} by the pressure only, and also by the thermodynamic relation $DS/Dt = 0$ obtained from (3') by putting $Q = 0$ and $\mathbf{h} = 0$ or its integral $S = \text{constant}$ in **homentropic flow**, which is governed by the **adiabatic law** $p \propto \rho^\gamma$, where γ denotes the ratio of specific heat at constant pressure to that at constant volume. In particular, for a liquid, the density variation can be neglected. Putting $\rho = \text{constant}$ in (1), we have

$$\operatorname{div} \mathbf{v} = 0, \quad (5)$$

which, in conjunction with (4), determines four unknowns (u, v, w, p) as functions of (x, y, z, t).

A fluid of constant density is called an **incompressible fluid**, and one of variable density a **compressible fluid**. Even though it might seem natural to consider gases as examples of compressible fluids, they can be treated as incompressible fluids if the speed of the flow of the gas $q = |\mathbf{v}|$ is small compared with the velocity $c = \sqrt{dp/d\rho}$ of sound propagating in the gas. We call $q/c = M$ the **Mach number**.

The vector $\boldsymbol{\omega}(\xi, \eta, \zeta)$, which is derived from the velocity vector \mathbf{v} as $\boldsymbol{\omega} = \operatorname{rot} \mathbf{v}$, is called the **vorticity**. A small part of the fluid rotates with angular velocity $\boldsymbol{\omega}/2$. If $\boldsymbol{\omega} = 0$, the flow is called **irrotational**, otherwise **rotational**. The curves $dx:dy:dz = u:v:w$ and $dx:dy:dz = \xi:\eta:\zeta$ are called, respectively, **stream lines** and **vortex lines**. The line integral $\oint_C \mathbf{v} \cdot d\mathbf{s}$ along a closed circuit C is called the **circulation** around C .

In irrotational flow, the velocity is expressed as $\mathbf{v} = \operatorname{grad} \Phi$, where Φ is called a **velocity potential**. When the external force \mathbf{K} has a potential Ω ($\mathbf{K} = -\operatorname{grad} \Omega$) and p is a definite function of ρ , we have the **pressure equation**

$$\frac{\partial \Phi}{\partial t} + \frac{1}{2} q^2 + \int \frac{dp}{\rho} + \Omega = \text{constant},$$

which is valid everywhere in the flow. In a steady flow,

$$\frac{1}{2} q^2 + \int \frac{dp}{\rho} + \Omega = \text{constant} \quad (6)$$

is valid along each stream line or each vortex line; this is called the **Bernoulli theorem**. These two equations correspond to energy integrals of the equation of motion. Furthermore, corresponding to the conservation of angular momentum, we have **Helmholtz's vorticity theorem**: When $\mathbf{K} = -\operatorname{grad} \Omega$ and $p = f(\rho)$, vorticity is neither created nor annihilated in the fluid.

For the irrotational motion of an incompressible fluid, $\nabla^2\Phi=0$ is derived from (5). Hence the problem reduces to the determination of a \dagger harmonic function Φ under appropriate boundary conditions (e.g., for a fixed wall, normal velocity $v_n=\partial\Phi/\partial n=0$). For the 2-dimensional problem a **stream function** Ψ is introduced to satisfy (5) by the relation $u=\partial\Psi/\partial y$, $v=-\partial\Psi/\partial x$. Since the \dagger Cauchy-Riemann equations $\partial\Phi/\partial x=\partial\Psi/\partial y$, $\partial\Phi/\partial y=-\partial\Psi/\partial x$ are valid in this case, $f=\Phi+i\Psi$ is an \dagger analytic function of $z=x+iy$. Therefore the theory of 2-dimensional irrotational motion is essentially equivalent to the theory of complex \dagger analytic functions, and consequently the theory of \dagger conformal mapping is a powerful method in the theory of such fluid motion.

For irrotational steady flow of a compressible fluid in which $\Omega=0$, c is determined from (6) as a function of q . Then (1) and (4) yield a \dagger nonlinear partial differential equation for Φ :

$$\left(1-\frac{u^2}{c^2}\right)\frac{\partial^2\Phi}{\partial x^2}+\left(1-\frac{v^2}{c^2}\right)\frac{\partial^2\Phi}{\partial y^2}+\left(1-\frac{w^2}{c^2}\right)\frac{\partial^2\Phi}{\partial z^2}-2\frac{vw}{c^2}\frac{\partial^2\Phi}{\partial y\partial z}-2\frac{wu}{c^2}\frac{\partial^2\Phi}{\partial z\partial x}-2\frac{uv}{c^2}\frac{\partial^2\Phi}{\partial x\partial y}=0. \quad (7)$$

This equation is \dagger elliptic or \dagger hyperbolic (\rightarrow 326 Partial Differential Equations of Mixed Type) according as M is less than 1 (**subsonic**) or greater than 1 (**supersonic**).

For 2-dimensional flow, we can introduce a stream function Ψ from (1) by $u=\partial\Phi/\partial x=(1/\rho)(\partial\Psi/\partial y)$, $v=\partial\Phi/\partial y=-(1/\rho)(\partial\Psi/\partial x)$. By utilizing the idea of \dagger Legendre transformation, this system of nonlinear equations for Φ and Ψ can be reduced to a system of linear equations in the **hodograph plane** (q, θ):

$$\frac{\partial\Phi}{\partial q}=q\frac{d}{dq}\left(\frac{1}{\rho q}\right)\frac{\partial\Psi}{\partial\theta}=-\frac{1-M^2}{\rho q}\frac{\partial\Psi}{\partial\theta},$$

$$\frac{\partial\Phi}{\partial\theta}=\frac{q}{\rho}\frac{\partial\Psi}{\partial q}$$

($d(\rho q)/dq=\rho(1-M^2)$), where the independent variables q and θ are the magnitude and the inclination of the velocity, respectively. The treatment of 2-dimensional compressible flow on the basis of this system is called the **hodograph method**. For a flow of small M , there is a method of successive approximation (M^2 -**expansion method**) which starts from Laplace's equation, neglecting the terms of $O(M^2)$ in (7). For uniform flow (velocity U in the x -direction) past a thin wing or slender body where v and w are small, we have **thin wing theory** or **slender body theory**, whose first approximation is

$$(1-M^2)\frac{\partial^2\Phi}{\partial x^2}+\frac{\partial^2\Phi}{\partial y^2}+\frac{\partial^2\Phi}{\partial z^2}=0. \quad (8)$$

If $M < 1$ or $M > 1$ (although not too large) a linearization (**Prandtl-Glauert approximation**) is possible by replacing M by the Mach number at infinity $M_\infty=U/c_\infty$. For $M > 1$, (8) has a \dagger characteristic surface, which is the **Mach cone** whose central axis makes an angle $\arcsin c/q = \arcsin 1/M$ with the flow. This can be interpreted also as an envelope produced by spherical sound waves with velocity c from a source drifting with velocity q . For $M \sim 1$, we put $\Phi=c_*x+\varphi$ (c_* is the fluid velocity when $q=c$). Then for an adiabatic gas, (8) is approximated by a partial differential equation of \dagger mixed type:

$$\frac{\partial^2\varphi}{\partial y^2}+\frac{\partial^2\varphi}{\partial z^2}=\frac{\gamma+1}{c_*}\frac{\partial\varphi}{\partial x}\frac{\partial^2\varphi}{\partial x^2}. \quad (9)$$

Such a flow in which both domains $M \gtrless 1$ co-exist is called the **transonic flow**, and exact solutions by the hodograph method are known. However, continuous deceleration from $M > 1$ to $M < 1$ generally tends to be unstable or impossible, and the appearance of a **shock wave**, i.e., a discontinuous surface of state quantities, is not unusual. This can be considered as the \dagger weak solution of (1), (2) (3) for a perfect fluid. In particular, in the coordinate system fixed to the surface, its integrated form can be obtained as follows: $[\rho v_n]=0$, $[p\delta_m+\rho v_i v_n]=0$, $[q^2/2+E+p/\rho]=0$ ($[]$ is the jump of the quantity at the surface, and n is the normal component). Supplemented by the entropy increase, these formulas give relations between the fluid velocity and the state variables at the front and back of the shock. In an ideal gas they are called the **Rankine-Hugoniot relation**. Entropy is not uniform behind a curved shock, and the flow is not irrotational. For a weak shock starting from the tip of a pointed slender body, however, the discontinuity is small and approaches the \dagger characteristic surface of (8), i.e., the **Mach wave** (compressive wave, in this case). Rarefactive Mach waves are found in the supersonic flow of acceleration around a convex surface. Such waves contribute to the drag on an obstacle placed in supersonic flow.

C. Viscous Fluids

A body moving uniformly in a fluid at rest (with velocity less than that of sound) suffers no drag as long as the viscosity of the fluid is negligible and the flow is continuous (**d'Alembert's paradox**). Hence we must take the viscosity into account in order to discuss the creation and annihilation of vortices, the generation and structure of shock waves, and the drag acting on obstacles. For this purpose, we extend **Newton's law** stating that frictional stress is proportional to the velocity gradient

and assume that the stress tensor \mathbf{p} is a linear function of the rate-of-strain tensor \mathbf{e} :

$$\begin{aligned} p_{xx} &= -p + \mu' \operatorname{div} \mathbf{v} + \mu e_{xx} - \frac{2}{3} \mu \operatorname{div} \mathbf{v} \\ &= -p + 2\mu \frac{\partial u}{\partial x} + \left(\mu' - \frac{2}{3} \mu \right) \operatorname{div} \mathbf{v}, \dots, \\ p_{yz} &= \mu e_{yz} = \mu \left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} \right), \dots \end{aligned}$$

The proportionality constants μ and μ' are called, respectively, the **coefficients of (shear) viscosity and bulk viscosity**. The bulk viscosity is sometimes neglected (**Stokes's assumption**) in the usual hydrodynamics, and then the mean value of the normal stress components equals the pressure. When a fluid satisfies this linear relation between \mathbf{p} and \mathbf{e} , it is called a **Newtonian fluid**. Otherwise, it is called a **non-Newtonian fluid**. Except for a few cases, such as colloid solutions, fluids can be regarded as Newtonian.

If we take the viscosity into account, the equation of motion of an incompressible fluid becomes

$$\rho D\mathbf{v}/Dt = \rho \mathbf{K} - \operatorname{grad} p + \mu \Delta \mathbf{v}. \quad (10)$$

This is called the **Navier-Stokes equation**. A nondimensional quantity $R = \rho UL/\mu$ formed by representative length L , velocity U , density ρ , and viscosity μ of a flow is called the **Reynolds number**. In order for two flows with geometrically similar boundaries to share similar kinetic properties, their Reynolds numbers must be equal. This is called the **Reynolds law of similarity**.

For small R , we can approximate the equation of motion (10) by replacing the acceleration $D\mathbf{v}/Dt$ by $\partial \mathbf{v}/\partial t$ (**Stokes approximation**) or by $\partial \mathbf{v}/\partial t + U \partial \mathbf{v}/\partial x$ (**Oseen approximation**) for a body placed in the uniform flow of velocity U in the x -direction.

For large R , the flow can be regarded as that of a perfect fluid, since we can neglect $\mu \Delta \mathbf{v}$ as long as the velocity gradient is not too large. In the vicinity of a fixed wall, however, the velocity gradient becomes large, because in a very thin layer the velocity decreases rapidly from the value U of a perfect fluid to zero at the wall. This layer is called the **boundary layer**. For the boundary layer, **Prandtl's boundary layer equation**

$$\begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} &= \frac{\partial U}{\partial t} + U \frac{\partial U}{\partial x} + \frac{\mu}{\rho} \frac{\partial^2 u}{\partial y^2}, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= 0 \end{aligned} \quad (11)$$

is valid, where x and y are the coordinates parallel and perpendicular to the wall, respec-

tively, and U is the velocity outside the boundary layer.

If $\partial U/\partial x < 0$, it sometimes happens that the boundary layer separates from the surface of the body. In this case a vortex is generated in the flow, as large vorticities in the boundary layer are carried into the flow. For a body without separation of the boundary layer, the d'Alembert paradox holds and is no longer a "paradox," and the drag is small. Such bodies are called **streamlined**.

In incompressible flow the new problem arises of the interaction of shock waves with the boundary layer. A rapid increase in pressure due to the shock wave formed on the surface of a body invalidates the assumption of a boundary layer and causes its separation. If the Mach number becomes sufficiently large ($M \gtrsim 5$, **hypersonic flow**), the bow shock approaches the body and interferes with the boundary layer. The generation of heat at the boundary layer (e.g., viscous dissipation in Q) requires the consideration of heat transfer as well as viscosity. In this manner, it becomes necessary to treat a complete system of equations which take into account the energy equation (3) as well as the temperature dependence of κ , μ , and μ' .

D. Laws of Similarity

For such complicated systems, \dagger dimensional analysis is often useful (\rightarrow 116 Dimensional Analysis). As laws of similarity, we can consider not only those like the Reynolds law but also others for bodies which transform similarly by \dagger affine transformations. Corresponding to equation (8), the **Prandtl-Glauert law of similarity** for subsonic flow is famous: The pressure coefficient (nondimensional pressure change) for a thin wing of chord (i.e., the length in the direction of flow) l , span L , and thickness τ is

$$C_p(L, \tau) = \lambda C_{p_0}(\sqrt{1 - M_\infty^2} L, \tau/\sqrt{1 - M_\infty^2} \lambda),$$

when λ is an arbitrary constant and C_{p_0} is C_p for a body of scaled length and thickness placed in an incompressible flow. Corresponding to (9), an extension of the famous **von Kármán transonic similarity** is possible:

$$\begin{aligned} C_p(L, \tau) &= \tau^{2/3}(\gamma + 1)^{-1/3} \\ &\times f(\sqrt{|1 - M_\infty^2|} L, (\gamma + 1)\tau/|1 - M_\infty^2|^{3/2}). \end{aligned}$$

E. Turbulence

For low Reynolds numbers, the flow generally has smooth streamlines. For high Reynolds numbers, however, extremely irregular motion in space and time appears. The former is called

laminar flow, and the latter **turbulent flow** (→ 433 Turbulence and Chaos). The transition from laminar to turbulent flow is considered to be due to the instability of the laminar flow, and this transition has been studied by the method of small oscillations. Recently, nonlinear effects have also been examined. Regarding the internal structure of turbulence, statistical theories originated by T. von Kármán and G. I. Taylor (*Proc. Roy. Soc. London*, 151 (1935)) and A. N. Kolmogorov (*Dokl. Akad. Nauk. SSSR*, 30 (1941)) are of central importance.

F. Water Waves

†Surface waves that occur on the free surface of water (or other liquids) are called **water waves**; their restoring forces are gravity and surface tension. If we consider waves generated on still water (assumed to be inviscid and incompressible) in equilibrium, we can regard the flow field associated with the wave motion as †irrotational by virtue of †Helmholtz's vorticity theorem. Hence the flow velocity can be derived from the †velocity potential Φ which satisfies †Laplace's equation $\Delta\Phi=0$, together with the boundary conditions

$$\frac{D(H+z)}{Dt} = \frac{\partial\Phi}{\partial x} \frac{\partial H}{\partial x} + \frac{\partial\Phi}{\partial y} \frac{\partial H}{\partial y} + \frac{\partial\Phi}{\partial z} = 0$$

at $z = -H(x, y)$, (12)

$$\frac{D(h-z)}{Dt} = \frac{\partial h}{\partial t} + \frac{\partial\Phi}{\partial x} \frac{\partial h}{\partial x} + \frac{\partial\Phi}{\partial y} \frac{\partial h}{\partial y} - \frac{\partial\Phi}{\partial z} = 0$$

at $z = z(x, y, t)$, (13)

$$\frac{\partial\Phi}{\partial t} + \frac{1}{2}(\text{grad } \Phi)^2 + gz = \frac{p_0 - p}{\rho} = \frac{\sigma}{\rho R_m}$$

at $z = h(x, y, t)$, (14)

where the Cartesian coordinates x and y are taken in the undisturbed horizontal free surface, while the positive z -axis is vertically upward. The equations $z = -H(x, y)$ and $z = h(x, y, t)$ denote, respectively, the bottom surface (assumed to be known) and the elevation of the free surface measured from the undisturbed level $z=0$, while g stands for the gravitational acceleration, σ the surface tension, p_0 the atmospheric pressure, and $1/R_m$ the †mean curvature of the disturbed free surface expressed as

$$\frac{1}{R_m} = \left[\left\{ 1 + \left(\frac{\partial h}{\partial x} \right)^2 \right\} \frac{\partial^2 h}{\partial y^2} + \left\{ 1 + \left(\frac{\partial h}{\partial y} \right)^2 \right\} \frac{\partial^2 h}{\partial x^2} - 2 \frac{\partial h}{\partial x} \frac{\partial h}{\partial y} \frac{\partial^2 h}{\partial x \partial y} \right] / \left\{ 1 + \left(\frac{\partial h}{\partial x} \right)^2 + \left(\frac{\partial h}{\partial y} \right)^2 \right\}^{3/2}$$

The conditions (12) and (13) imply, respectively, that the fluid does not cross the bottom and the free surface, while (14) expresses the fact that the difference between atmospheric and fluid pressures at the free surface is equal to the normal force (per unit area) due to the surface tension. Thus the problem is formulated as a nonlinear boundary value problem for Laplace's equation including the unknown boundary $z = h(x, y, t)$.

Let us first consider linear waves for which the wave amplitude of the surface elevation is much smaller than any other characteristic linear dimension such as the wavelength or the water depth H (for simplicity, we assume hereafter $H = \text{constant}$, i.e., a flat horizontal bottom). Linearizing the boundary conditions (12)–(14) with respect to h and $\text{grad } \Phi$, and assuming a sinusoidal wave proportional to $\exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)]$, $\mathbf{k}(k_x, k_y)$ and ω being respectively the (2-dimensional) †wave number vector and the †angular frequency and $\mathbf{r} = (x, y)$, we obtain the dispersion relation

$$\omega^2 = \left(1 + \frac{\sigma k^2}{\rho g} \right) g k \tanh(kH). \quad (15)$$

In a layer of still water the waves are isotropic in the horizontal plane and the dispersion relation involves only the magnitude k of the wave number vector. It is readily seen from (15) that the water waves are typical †dispersive waves in which the †phase velocity $c_p (= \omega/k)$ depends on the wave number k or the wavelength $\lambda (= 2\pi/k)$. It is also evident from (15) that the quantity $\sigma k^2/(\rho g)$ measures the relative importance of surface tension and gravity. Hence for waves with wavelengths much larger than $\lambda_m = 2\pi\sqrt{\sigma/(\rho g)}$ (~ 1.7 cm for water), the effect of surface tension is negligible, and we have **gravity waves**. Conversely, when $\lambda \ll \lambda_m$, the effect of surface tension becomes dominant, and we have **capillary waves** or **ripples**. When the water depth H is much larger than the wavelength λ , we can approximate (15) by $\omega^2 = gk + \sigma k^3/\rho$, since $\tanh(kH) \sim 1$. We call such waves **deep water waves**. On the other hand, if H is much smaller than the wavelength ($kH \ll 2\pi$), we have **shallow** (or **long**) **water waves** for which (15) can be approximated by $\omega^2 = gHk^2 [1 + \{\sigma/(\rho gH^2) - 1/3\} (kH)^2 + \dots]$ if $\sigma/(\rho gH^2) = O(1)$. In particular, if we neglect $O(kH)^2$, we recover the well-known dispersionless **long gravity waves** whose phase velocity is simply \sqrt{gH} . In all the cases mentioned above, the amplitude function of the velocity potential Φ is proportional to $-i\omega \cosh\{k(z+H)\}/\{k \sinh(kH)\}$, so that the flow velocity due to deep water waves decreases exponentially as one proceeds vertically downward from the free surface. In the

limiting case of long gravity waves, however, the fluid motion is nearly horizontal throughout the fluid layer.

When the wave amplitude becomes larger, nonlinear effects are no longer negligible. For such waves, various †singular perturbation methods provide powerful tools. For example, the basic system of equations for weakly nonlinear (1-dimensional) shallow water waves can be reduced to a simple solvable nonlinear equation called the †Korteweg-de Vries equation, whose solitary wave solution is known as a prototype of solitons (\rightarrow 387 Solitons). Another classical example is a (1-dimensional) deep gravity wave called the **Stokes wave**, which can be obtained as a power series in the **wave steepness** (amplitude \times wave number). The first term of the series is of the form of a linear sinusoidal wave, and the higher-order terms correspond to the higher harmonics, while the angular frequency is shifted from the linear case and depends not only on the wave number but also on the amplitude. Similar singular perturbation methods have also been applied to various kinds of resonant interactions such as nonlinear self-modulation, higher harmonic resonances, and multiwave interactions. Finally it should be mentioned that an exact solution representing a (1-dimensional) deep capillary wave was obtained by G. D. Crapper (*J. Fluid Mech.*, 2 (1957) 532–540; extended later to the case of finite depth by W. Kinnersley, *J. Fluid Mech.*, 77 (1976), 229–241). This is one of the few realistic exact solutions obtained so far; a famous exact solution of Gerstner's trochoidal wave does not satisfy the irrotationality condition.

References

- [1] H. Lamb, *Hydrodynamics*, Cambridge Univ. Press, sixth edition, 1932.
- [2] S. Goldstein, *Modern developments in fluid dynamics*, Clarendon Press, 1938.
- [3] R. Courant and K. O. Friedrichs, *Supersonic flow and shock waves*, Interscience, 1948.
- [4] L. Howarth, *Modern developments in fluid dynamics*, High speed flow I, II, Oxford Univ. Press, 1953.
- [5] K. Oswatitsch, *Gas dynamics*, Academic Press, 1956.
- [6] H. W. Liepmann and A. Roshko, *Elements of gasdynamics*, Wiley, 1956.
- [7] J. J. Stoker, *Water waves*, Interscience, 1957.
- [8] L. D. Landau and E. M. Lifshitz, *Fluid mechanics*, Pergamon, 1959. (Original in Russian, 1954.)
- [9] K. G. Guderley, *The theory of transonic flow*, Pergamon, 1962.

[10] L. Rosenhead, *Laminar boundary layers*, Clarendon Press, 1963.

[11] G. K. Batchelor, *An introduction to fluid dynamics*, Cambridge Univ. Press, 1967.

[12] G. B. Whitham, *Linear and nonlinear waves*, Wiley, 1974.

[13] M. J. Lighthill, *Waves in fluids*, Cambridge Univ. Press, 1978.

[14] H. Schlichting, *Boundary-layer theory*, McGraw-Hill, 1979.

206 (XIV.5) Hypergeometric Functions

A. Hypergeometric Functions

The †power series

$$F(\alpha, \beta, \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\beta)} \sum_{n=1}^{\infty} \frac{\Gamma(\alpha+n)\Gamma(\beta+n)}{n!\Gamma(\gamma+n)} z^n$$

in the complex variable z is called the **hypergeometric series** or **Gauss's series**.

It is convergent for any α , β , and γ if $|z| < 1$, and is convergent for $\text{Re}(\alpha + \beta - \gamma) < 0$ if $|z| = 1$. If $z = 1$, its sum is equal to $\Gamma(\gamma)\Gamma(\gamma - \alpha - \beta)/\Gamma(\gamma - \alpha)\Gamma(\gamma - \beta)$ (except when γ is a nonpositive integer). The **hypergeometric functions** are obtained as analytic continuations of the functions determined by hypergeometric series that are single-valued analytic functions defined on the domain obtained from the complex plane by deleting a line connecting branch points $z = 1$ and $z = \infty$ (\rightarrow Appendix A, Table 18.1).

A hypergeometric function is a solution of the differential equation

$$z(1-z)\frac{d^2w}{dz^2} + (\gamma - (\alpha + \beta + 1)z)\frac{dw}{dz} - \alpha\beta w = 0, \quad (1)$$

which is called the **hypergeometric differential equation** or **Gaussian differential equation**. This equation is a differential equation of †Fuchsian type with †regular singular points at 0, 1, and ∞ , whose solutions are expressed, in terms of the † P -function of Riemann, by

$$w = P \left\{ \begin{matrix} 0 & \infty & 1 \\ 0 & \alpha & 0 \\ 1-\gamma & \beta & \gamma-\alpha-\beta \end{matrix} \right\} z.$$

If any one of the values of γ , $\gamma - \alpha - \beta$, or $\alpha - \beta$ is integral, there exists a series containing $\log z$, representing a solution of the differential equation (1) in a neighborhood of the corresponding singular points. When none of the γ , $\gamma - \alpha - \beta$, or $\alpha - \beta$ values is integral, since the linear transformations $z' = z$, $z' = 1/z$, $z' = 1 - z$,

$z' = z/(z-1)$, $z' = (z-1)/z$, $z' = 1/(1-z)$ permute singular points, there exist 24 particular solutions around the singular points. The latter fact was first proved by E. E. Kummer (1836).

There exist various curves C for which the integral

$$w = \int_C u^{\alpha-1} (1-u)^{\gamma-\alpha-1} (1-zu)^{-\beta} du$$

is a solution of (1). Among them we can take the segment $[0, 1]$ when $\operatorname{Re} \alpha > 0$, $\operatorname{Re}(\gamma - \alpha) > 0$. Then the corresponding solution is holomorphic in the interior of the unit circle, and

$$F(\alpha, \beta, \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \times \int_C u^{\alpha-1} (1-u)^{\gamma-\alpha-1} (1-zu)^{-\beta} du.$$

Since the integrand has branch points at 0, 1, and $1/z$, we have the following expression when γ is not an integer:

$$F(\alpha, \beta, \gamma; z) = \frac{1}{(1-e^{2\pi i(\gamma-\alpha)})(1-e^{2\pi iz})} \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \times \oint^{(1+, 0+, 1-, 0-)} u^{\alpha} (1-u)^{\gamma-\alpha-1} (1-zu)^{-\beta} du,$$

where $\operatorname{Re} \alpha > 0$, $\operatorname{Re}(\gamma - \alpha) > 0$; whereas if γ is an integer, then

$$F(\alpha, \beta, \gamma; z) = \frac{1}{(1-e^{-2\pi iz})} \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \times \oint^{(1+, 0+)} u^{\alpha} (1-u)^{\gamma-\alpha-1} (1-zu)^{-\beta} du,$$

where the contour in the first expression encircles successively each of 1, 0, 1, and 0 once with indicated directions. These expressions can be adopted as a definition of the hypergeometric functions for the general value of z . Other integral expressions are also known (\rightarrow Appendix A, Table 18.I).

B. The Ladder Method

A linear ordinary differential equation of the second order having three regular singular points on the complex sphere is easily transformed into an equation of the form (1). To solve such an equation with a parameter, it is often useful to decompose, in two different ways, the main part of the equation into two factors of the first order, and find a recurrence formula involving the parameter, as we shall see in the following example. This method is called the **ladder method** or **factorization method**.

For example, [†]Legendre's differential

equation

$$L_n[w] \equiv (1-z^2)((1-z^2)w)' + n(n+1)w = 0$$

is decomposed as follows:

$$L_n = S_n \cdot T_n + n^2 = T_{n+1} \cdot S_{n+1} + (n+1)^2,$$

$$T_n = (1-z^2) \frac{d}{dz} + nz, \quad S_n = (1-z^2) \frac{d}{dz} - nz.$$

If w_n is a solution of $L_n[w] = 0$, then multiplying both sides of $S_n \cdot T_n[w_n] + n^2 w_n = 0$ by T_n , we find that $T_n \cdot S_n(T_n[w_n]) + n^2(T_n[w_n]) = 0$, that is, $T_n[w_n]$ is a solution of $L_{n-1}[w] = 0$. Similarly, we see that $S_{n+1}[w_n]$ is a solution of $L_{n+1}[w] = 0$. In this sense, S_n and T_n are called, respectively, the **step-up operator**, or **up-ladder**, and the **step-down operator**, or **down-ladder**, with respect to the parameter n .

The above relation constitutes a recurrence formula for Legendre functions (\rightarrow Appendix A, Table 18.II).

C. Extensions of Hypergeometric Functions

J. Thomas (1870) proposed the series

$$1 + \sum_{n=1}^{\infty} \frac{(\alpha_1)_n (\alpha_2)_n \dots (\alpha_h)_n}{(\beta_1)_n (\beta_2)_n \dots (\beta_h)_n} z^n, \\ (\lambda)_n = \lambda(\lambda+1) \dots (\lambda+n-1)$$

as an extension of the hypergeometric series. The sum of this series satisfies the h th-order differential equation

$$(1-z) \frac{d^h w}{dt^h} + (A_1 - B_1 z) \frac{d^{h-1} w}{dt^{h-1}} \\ + (A_2 - B_2 z) \frac{d^{h-2} w}{dt^{h-2}} + \dots + (A_h - B_h z) w = 0,$$

$$t = \log z.$$

When $h=2$ and $\beta_1=1$, it reduces to the ordinary hypergeometric series. The notation

$${}_p F_q(\alpha_1, \alpha_2, \dots, \alpha_p; \beta_1, \beta_2, \dots, \beta_q; z) \\ = \sum_{n=0}^{\infty} \frac{(\alpha_1)_n (\alpha_2)_n \dots (\alpha_p)_n}{n! (\beta_1)_n (\beta_2)_n \dots (\beta_q)_n} z^n, \quad (2)$$

which is due to L. Pochhammer and modified by E. W. Barnes, is used to denote the extended hypergeometric series, and the function defined by (2) is often called **Barnes's extended hypergeometric function**. For example, Gauss's series in this notation is ${}_2F_1(\alpha, \beta, \gamma; z)$.

Corresponding to Barnes's integral expression for hypergeometric functions, it is known that the integral

$$W(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} K(\zeta) H(\zeta) z^{-\zeta} d\zeta,$$

where

$$K(\zeta) = K(\zeta+1)$$

and

$$H(\zeta) = \frac{\Gamma(\zeta + \alpha_1)\Gamma(\zeta + \alpha_2) \dots \Gamma(\zeta + \alpha_h)}{\Gamma(\zeta + 1 + \beta_1)\Gamma(\zeta + 1 + \beta_2) \dots \Gamma(\zeta + 1 + \beta_h)},$$

is a solution of the h th-order differential equation at the beginning of this section. The hypergeometric function expressed by the definite integral

$$\int \zeta^a (\zeta - 1)^b (\zeta - z)^c d\zeta$$

has an obvious formal extension

$$\int (\zeta - a_1)^{b_1} (\zeta - a_2)^{b_2} \dots (\zeta - a_m)^{b_m} (\zeta - z)^c d\zeta.$$

On the other hand, the equation

$$\sum_{v=0}^m \varphi_v(z) \frac{d^v w}{dz^v} = 0,$$

where

$$\varphi_v(z) = \frac{(-1)^{m-1-v}}{(h+m-2) \dots (h+1)h} \times \left(\binom{h+m-v-2}{m-v-1} P_1^{(m-1-v)}(z) + \binom{h+m-v-2}{m-v} P_0^{(m-v)}(z) \right),$$

$$P_0(z) = (z - a_1)(z - a_2) \dots (z - a_m),$$

$$P_1(z) = P_0(z) \left(\frac{\beta_1}{z - a_1} + \frac{\beta_2}{z - a_2} + \dots + \frac{\beta_m}{z - a_m} \right),$$

called the **Tissot-Pochhammer differential equation**, has a solution

$$w(z) = \int_c (\zeta - a_1)^{\beta_1-1} (\zeta - a_2)^{\beta_2-1} \dots \times (\zeta - a_m)^{\beta_m-1} (\zeta - z)^{h+m-2} d\zeta.$$

After Pochhammer (1870), this is sometimes called Pochhammer's generalized hypergeometric function.

As another extension of Gauss's series, H. E. Heine (1846) introduced **Heine's series**:

$$\varphi(a, b, c; q; z) = 1 + \frac{(1-q^a)(1-q^b)}{(1-q)(1-q^c)} q^z + \frac{(1-q^a)(1-q^{a+1})(1-q^b)(1-q^{b+1})}{(1-q)(1-q^2)(1-q^c)(1-q^{c+1})} q^{2z} + \dots$$

Setting $q = 1 + \varepsilon$, $z = (1/\varepsilon) \log x$, and letting $\varepsilon \rightarrow 0$, we obtain Gauss's series as the limit of Heine's series.

D. Hypergeometric Functions of Several Variables

P. Appell (1880) formally extended Gauss's series to the case of two variables and defined

Hypergeometric Functions

four kinds of functions [3]:

$$F_1(\alpha; \beta, \beta'; \gamma; x, y)$$

$$= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\alpha)_{m+n} (\beta)_m (\beta')_n}{m! n! (\gamma)_{m+n}} x^m y^n,$$

$$F_2(\alpha; \beta, \beta'; \gamma, \gamma'; x, y)$$

$$= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\alpha)_{m+n} (\beta)_m (\beta')_n}{m! n! (\gamma)_m (\gamma')_n} x^m y^n,$$

$$F_3(\alpha, \alpha'; \beta, \beta'; \gamma; x, y)$$

$$= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\alpha)_m (\alpha')_n (\beta)_m (\beta')_n}{m! n! (\gamma)_{m+n}} x^m y^n,$$

$$F_4(\alpha; \beta; \gamma, \gamma'; x, y)$$

$$= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\alpha)_{m+n} (\beta)_{m+n}}{m! n! (\gamma)_m (\gamma')_n} x^m y^n.$$

They are called **Appell's hypergeometric functions of two variables**. Each satisfies a corresponding system of partial differential equations:

$$F_1 \begin{cases} x(1-x)r + y(1-x)s + (\gamma - cx)p - \beta yq \\ \quad \quad \quad - \alpha \beta z = 0, \\ y(1-y)t + x(1-y)s + (\gamma - c'x)q - \beta' xp \\ \quad \quad \quad - \alpha \beta' z = 0, \\ (x-y)s - \beta' p + \beta q = 0, \end{cases}$$

$$F_2 \begin{cases} x(1-x)r - xys + (\gamma - cx)p - \beta yq - \alpha \beta z = 0, \\ y(1-y)t - xys + (\gamma' - c'y)q - \beta' xp \\ \quad \quad \quad - \alpha \beta' z = 0, \end{cases}$$

$$F_3 \begin{cases} x(1-x)r + ys + (\gamma - cx)p - \alpha \beta z = 0, \\ y(1-y)t + xs + (\gamma - c''y)q - \alpha' \beta' z = 0, \end{cases}$$

$$F_4 \begin{cases} x(1-x)r - y^2 t - 2xys + (\gamma - cx)p - cyq \\ \quad \quad \quad - \alpha \beta z = 0, \\ y(1-y)t - x^2 r - 2xys + (\gamma' - cy)q - cxp \\ \quad \quad \quad - \alpha \beta' z = 0, \end{cases}$$

where

$$c = \alpha + \beta + 1, \quad c' = \alpha + \beta' + 1, \quad c'' = \alpha' + \beta' + 1, \\ p = \partial z / \partial x, \quad q = \partial z / \partial y, \quad r = \partial^2 z / \partial x^2, \\ s = \partial^2 z / \partial x \partial y, \quad t = \partial^2 z / \partial y^2.$$

Appell's hypergeometric functions can also be represented by integrals: for example,

$$F_1 = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\beta')\Gamma(\gamma - \beta - \beta')} \iint u^{\beta-1} v^{\beta'-1} \times (1-u-v)^{\gamma-\beta-\beta'-1} (1-ux-vy)^{-\alpha} du dv,$$

$$F_2 = \frac{\Gamma(\gamma)\Gamma(\gamma')}{\Gamma(\beta)\Gamma(\beta')\Gamma(\gamma - \beta)\Gamma(\gamma' - \beta')} \int_0^1 \int_0^1 u^{\beta-1} v^{\beta'-1} \times (1-u)^{\gamma-\beta-1} (1-v)^{\gamma'-\beta'-1} (1-ux-vy)^{-\alpha} du dv,$$

$$F_3 = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\beta')\Gamma(\gamma-\beta-\beta')} \iint u^{\beta-1} v^{\beta'-1} \times (1-u-v)^{\gamma-\beta-\beta'-1} (1-ux)^{-\alpha} (1-vy)^{-\alpha'} du dv,$$

where, for F_1 and F_3 , the domain of integration is $u \geq 0, v \geq 0, 1-u-v \geq 0$. E. Picard (1881) showed that F_1 can also be expressed by a single integral:

$$F_1 = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\alpha-\gamma)} \int_0^1 u^{\alpha-1} (1-u)^{\gamma-\alpha-1} \times (1-ux)^{-\beta} (1-uy)^{-\beta'} du.$$

G. Lauricella (1893) extended the foregoing functions to the case of more than two variables. More general hypergeometric series of several variables were defined by R. Mellin, J. Horn, and J. Kampé de Fériet [3, 4]. Every algebraic equation can be solved analytically in terms of the foregoing functions (Mellin, R. Birkeland) [4]. Also there are studies concerning [†]Riemann's problem and [†]automorphic functions derived from F_1 (Picard; T. Terada [5]).

E. Hypergeometric Functions with Matrix Argument

For symmetric matrices Z of degree m , C. S. Herz defined **hypergeometric functions with matrix argument** as follows [5]: Denoting by $\text{etr } Z$ the exponential $\exp(\text{tr } Z)$ of the [†]trace of Z , let

$$\begin{aligned} {}_0F_0(Z) &= \text{etr } Z, \\ {}_{p+1}F_q(\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q; \gamma; Z) \\ &= \frac{1}{\Gamma_m(\gamma)} \int_{\Lambda > 0} \text{etr}(-\Lambda) {}_pF_q(\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q; \\ &\quad \Lambda Z)(\det \Lambda)^{\gamma-p} d\lambda_{11} d\lambda_{22} \dots d\lambda_{mm}, \quad (3) \end{aligned}$$

$$\begin{aligned} {}_pF_{q+1}(\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q; \gamma; \Lambda) \\ &= \frac{\Gamma_m(\gamma)}{(2\pi i)^{m(m+1)/2}} \int_{\text{Re } Z = X_0 > 0} \text{etr } Z {}_pF_q(\alpha_1, \dots, \alpha_p; \\ &\quad \beta_1, \dots, \beta_q; \Lambda Z^{-1})(\det Z)^{-\gamma} dz_{11} dz_{22} \dots dz_{mm}, \quad (4) \end{aligned}$$

where

$$\Lambda = (\lambda_{ij})_{i,j=1,\dots,m},$$

$$Z = ((1 + \delta_{ij})z_{ij}/2)_{i,j=1,\dots,m},$$

$$\Gamma_m(\gamma) = \pi^{m(m-1)/4} \Gamma(\gamma) \Gamma(\gamma - 1/2) \dots$$

$$\times \Gamma(\gamma - (m-1)/2),$$

and $\Lambda > 0$ means that Λ is [†]positive definite. The integral (3) converges for $-Z > 0$ if $\text{Re } \gamma > (m-1)/2$. If $\text{Re } \gamma$ is sufficiently large, then for suitably chosen X_0 , (4) converges in a domain of the space of Λ and represents an analytic function of its argument. In particular, we

have

$${}_1F_0(x; Z) = (\det(E - Z))^{-x}.$$

Based on this definition, many special functions and formulas are extended to the case of a matrix argument. For example,

$$A_\delta(Z) = {}_0F_1(\delta + (m+1)/2; -Z)/\Gamma_m(\delta + (m+1)/2) \quad (5)$$

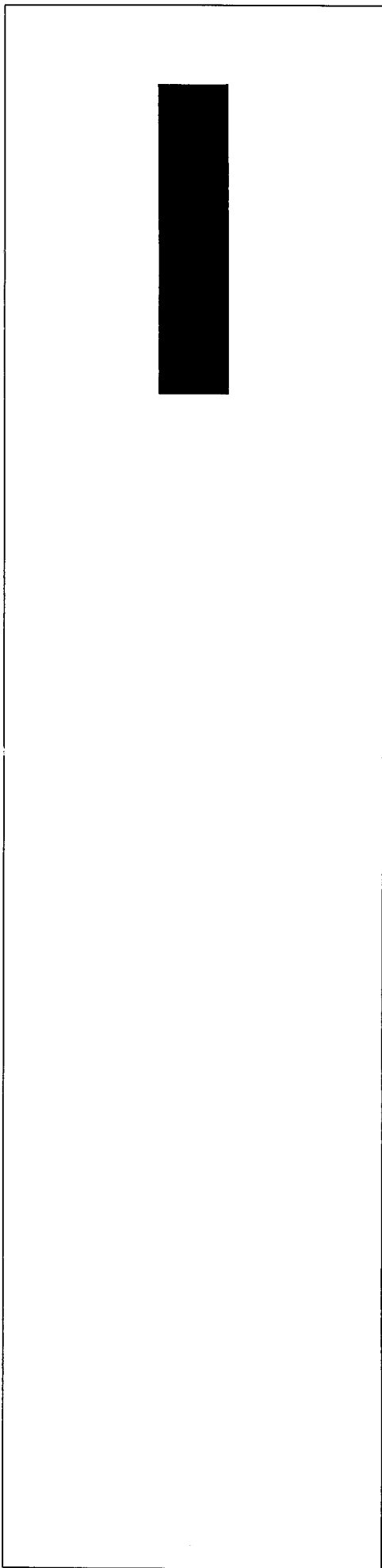
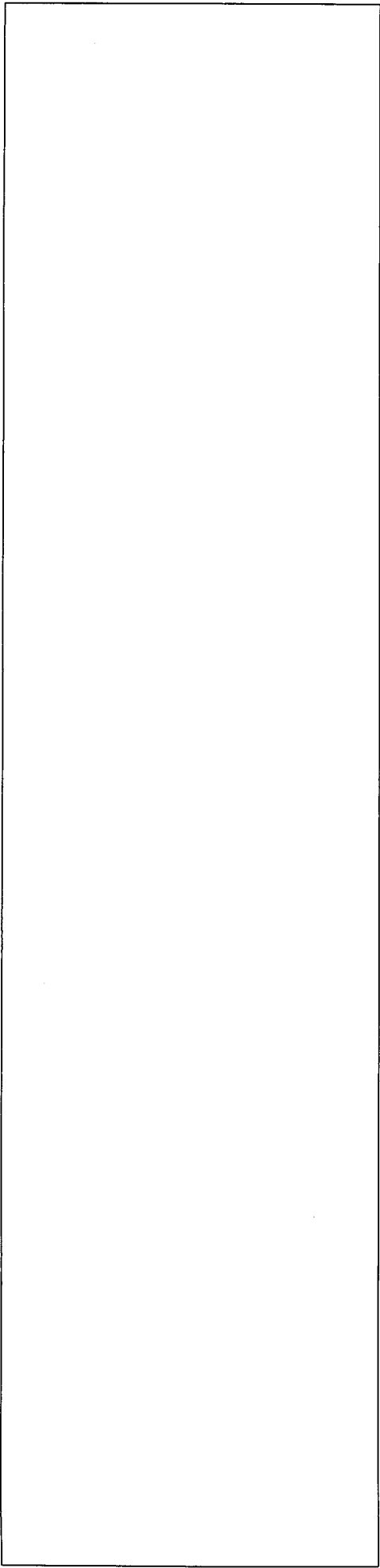
is an extension of the [†]Bessel function, and this reduces to

$$(t/2)^{-\delta} J_\delta(t) = A_\delta((t/2)^2)$$

when $m = 1$. Formula (5) is applied to the [†]noncentral Wishart distribution in mathematical statistics.

References

- [1] F. Klein, *Vorlesungen über die hypergeometrische Funktionen*, Springer, 1933. Also \rightarrow references to 389 Special Functions. For applications of the ladder method,
- [2] L. Infeld and T. E. Hull, The factorization method, *Rev. Mod. Phys.*, 23 (1951), 21-68. For hypergeometric functions of several variables,
- [3] P. Appell, *Sur les fonctions hypergéométriques de plusieurs variables*, *Mémor. Sci. Math.*, Gauthier-Villars, 1925.
- [4] G. Belardinelli, *Fonctions hypergéométriques de plusieurs variables et résolution analytique des équations algébriques générales*, *Mémor. Sci. Math.*, Gauthier-Villars, 1960.
- [5] T. Terada, *Problème de Riemann et fonctions automorphes provenant des fonctions hypergéométriques de plusieurs variables*, *J. Math. Kyoto Univ.* 13 (1973), 557-578. For the case of a matrix variable,
- [6] C. S. Herz, Bessel functions of matrix argument, *Ann. Math.*, (2) 61 (1955), 474-523.
- [7] L. J. Slater, *Generalized hypergeometric functions*, Cambridge Univ. Press, 1966.



207 (XI.13) Ideal Boundaries

A. Ideal Boundaries

For a given Hausdorff space R , a \dagger compact Hausdorff space R^* that contains R as its dense subspace is called a **compactification** of R , and $\Delta = R^* - R$ is called an **ideal boundary** of R . In the present article, we deal mainly with properties (in particular, function-theoretic properties) of ideal boundaries of \dagger Riemann surfaces R .

B. Harmonic Boundaries

By R we mean an open Riemann surface. The set Γ of points p^* in Δ such that $\liminf_{R \ni p \rightarrow p^*} P(p) = 0$ for every \dagger potential P (i.e., a positive \dagger superharmonic function P for which the class of nonnegative \dagger harmonic functions smaller than P consists only of the constant function 0) is a compact subset of R^* . The set Γ is called the **harmonic boundary** of R with respect to R^* . For an arbitrary compact subset K in $\Delta - \Gamma$, there exists a finite-valued potential P_K with $\lim_{R \ni p \rightarrow p^*} P_K(p) = \infty$ ($p^* \in K$). From this, various kinds of \dagger maximum principles are derived. For instance, if u is a harmonic function bounded above for which $\limsup_{p \rightarrow \Gamma} u(p) \leq M$ holds, then $u \leq M$ on R .

There are infinitely many compactifications of R . For two compactifications R_i^* ($i = 1, 2$) of R , we say that R_1^* is **greater than** R_2^* or, equivalently, **lies over** R_2^* , if the identity mapping of R can be extended to a continuous mapping of R_1^* onto R_2^* . In order that deep function-theoretic studies of R^* can be carried out, various conditions must be imposed on R^* . A compactification R^* is said to be of **Stoilow type** (or of **type S**) if for every \dagger connected open subset G^* in R^* whose boundary in R^* is contained in R , $G^* - \Delta$ is also connected. Next suppose that $R \notin O_G$ (\rightarrow 367 Riemann Surfaces E). For a given real-valued function f on Δ , let \underline{H}_f^{R, R^*} (\underline{H}_f^{R, R^*}) be the class of \dagger superharmonic functions s bounded from below (\dagger subharmonic functions s bounded from above) such that $\liminf_{R \ni p \rightarrow p^*} s(p) \geq f(p^*)$ ($\limsup_{R \ni p \rightarrow p^*} s(p) \leq f(p^*)$) for every $p^* \in \Delta$. If these classes are nonempty, then $\bar{H}_f^{R, R^*}(p) = \inf\{s(p) | s \in \underline{H}_f^{R, R^*}\}$ and $\underline{H}_f^{R, R^*}(p) = \sup\{s(p) | s \in \underline{H}_f^{R, R^*}\}$ are harmonic on R , and $\underline{H}_f^{R, R^*} \leq \bar{H}_f^{R, R^*}$. In particular, if $\underline{H}_f^{R, R^*} \equiv \bar{H}_f^{R, R^*}$, then the common function is denoted by H_f^{R, R^*} , and the function f is said to be **resolutive** with respect to R^* . A compactification such that every bounded continuous function on Δ is resolutive is called a **resolutive compactification**. In

such a case, a point p^* in Δ is said to be **regular** with respect to the \dagger Dirichlet problem if $\lim_{R \ni p \rightarrow p^*} H_f^{R, R^*}(p) = f(p^*)$ for every bounded continuous function f on Δ (\rightarrow 120 Dirichlet Problem). The set Δ_r of regular points in Δ is contained in Γ . If R^* is a resolutive compactification, then there exists a unique positive \dagger Borel measure μ_p such that $H_f^{R, R^*}(p) = \int_{\Delta} f(p^*) d\mu_p(p^*)$ for every bounded continuous function f on Δ . This measure is called the **harmonic measure** with respect to $p \in R$. There exists a function $P(p, p^*)$ on $R \times \Delta$ with $d\mu_p(p^*) = P(p, p^*) d\mu_o(p^*)$ for an arbitrary fixed point o in R satisfying the following three conditions: (i) $P(p, p^*)$ is harmonic on R as a function of p ; (ii) $P(p, p^*)$ is Borel measurable as a function of p^* ; (iii) $k(o, p)^{-1} \leq P(p, p^*) \leq k(o, p)$, with the Harnack constant $k(o, p)$ of $\{o, p\}$ relative to R [9].

C. Compactifications Determined by Function Families

A family F of real-valued continuous functions on R admitting infinite values is called a **separating family** on R if there exists an f in F such that $f(p) \neq f(q)$ for any pair of given distinct points p and q in R . A compactification R^* is called an **F -compactification**, denoted by R_F^* , if every function in F can be continuously extended to R^* and the family of extended functions again constitutes a separating family on R^* . The correspondence $\varphi: F \rightarrow R_F^*$ defines a single-valued mapping of all separating families F on R onto all F -compactifications of R . If $F_1 \supset F_2$, then $\varphi(F_1)$ lies over $\varphi(F_2)$. For any R^* , $\varphi^{-1}(R^*)$ contains infinitely many separating families, among which the separating families constituting \dagger associative algebras are important. The following are typical examples of compactifications determined by function families:

(1) The **Aleksandrov compactification** is the \mathcal{U} -compactification $R_{\mathcal{U}}^*$ with the family \mathcal{U} of bounded continuous functions on R with compact support. It is the smallest compactification of R and is often used in function theory in discussing Dirichlet problems for relatively noncompact subregions in reference to relative boundaries.

(2) The **Stone-Ćech compactification** is the \mathcal{C} -compactification $R_{\mathcal{C}}^*$ with the family \mathcal{C} of bounded continuous functions on R . It is the largest compactification of R . It is rarely used in function theory, but an application is found in the work of M. Nakai [8].

(3) The **Kerékjártó-Stoilow compactification** is the \mathcal{S} -compactification $R_{\mathcal{S}}^*$ with the family \mathcal{S} of bounded continuous functions f on R

such that there exist compact sets K_f with the property that the f are constants on each connected component of $R - K_f$. This is the smallest compactification of Stoilow type. Many applications of this compactification can be found in function theory, among which the investigation done by M. Ohtsuka on the Dirichlet problem and the theory of conformal mappings is typical.

(4) The **Royden compactification** is the \mathfrak{R} -compactification $R_{\mathfrak{R}}^*$ with the family \mathfrak{R} of bounded C^∞ functions f on R with finite Dirichlet integrals $\iint_R df \wedge *df$. It was introduced by H. L. Royden and developed further by S. Mori, M. Ôta, Y. Kusunoki, Nakai, and others. This compactification has been used effectively in the study of HD -functions and the classification problem of Riemann surfaces (\rightarrow 367 Riemann Surfaces).

(5) The **Wiener compactification** is the \mathfrak{W} -compactification $R_{\mathfrak{W}}^*$ with the family \mathfrak{W} of bounded continuous functions f on R such that $\{H_f^{G_n}\}$ converges to a unique harmonic function independent of the choice of exhaustions $\{G_n\}$ of an arbitrary fixed subregion $G \notin O_G$, where the G_n are relatively compact subregions of G . It is the largest resolutive compactification, and compactifications smaller than $R_{\mathfrak{W}}^*$ are always resolutive. This compactification was introduced independently by Mori, K. Hayashi, Kusunoki, and C. Constantinescu and A. Cornea and is useful for the study of HB -functions and the classification of Riemann surfaces.

(6) The **Martin compactification** is the \mathfrak{M} -compactification $R_{\mathfrak{M}}^*$ with the family \mathfrak{M} of bounded continuous functions f on R such that there exist relatively compact regions R_f with the property that $f = H_{f^*}^{R - R_f, R_{\mathfrak{M}}^* - R_f} / H_{1^*}^{R - R_f, R_{\mathfrak{M}}^* - R_f}$ on $R - R_f$. Here f^* coincides with f on R_f and equals 0 on $R_{\mathfrak{M}}^* - R$, and 1^* is similarly defined. The set $R_{\mathfrak{M}}^* - R$ is called the **Martin boundary** of R . If \dagger Green's function g exists on R , then the function $m(p, q) = g(p, q)/g(o, q)$ for an arbitrary fixed $o \in R$ can be extended continuously to $R \times R_{\mathfrak{M}}^*$, which is called the **Martin kernel**. By the metric $d_{\mathfrak{M}}(q, r) = \sup_{p \in R_0} |m(p, q)/(1 + m(p, q)) - m(p, r)/(1 + m(p, r))|$ with a parametric disk R_0 in R , $R_{\mathfrak{M}}^*$ is \dagger metrizable. This compactification was introduced by R. S. Martin, and many applications of it to the study of HP -functions, potential theory, Markov chains, and cluster sets were obtained by M. H. Heins, Z. Kuramochi, J. L. Doob, Constantinescu and Cornea, and others.

(7) For a function f on R , $(R)\partial f/\partial n = 0$ means that there exists a relatively compact subregion R_f such that f is of class C^∞ on R outside R_f and the Dirichlet integral of f over $R - R_f$ is not greater than those of functions

on $R - R_f$ that coincide with f on the boundary of R_f . The **Kuramochi compactification** is the \mathfrak{K} -compactification $R_{\mathfrak{K}}^*$ with the family \mathfrak{K} of bounded continuous functions f on R satisfying $(R)\partial f/\partial n = 0$. The continuous function $k(p, q)$ on R such that $(R)\partial k/\partial n = 0$ vanishes in a fixed parametric disk R_0 in R and is harmonic in $R - R_0$ except for a positive \dagger logarithmic singularity at a point q can be extended continuously to $R \times R_{\mathfrak{K}}^*$, which is called the **Kuramochi kernel**. By the use of this kernel, $R_{\mathfrak{K}}^*$ is metrizable, as in the case of Martin compactification. This compactification was introduced by Kuramochi, and its important applications to the study of HD -functions, potential theory, and cluster sets were made by Kuramochi, Constantinescu and Cornea, and others.

Among compactifications (1)–(7), no boundary point in (2), (4), or (5) satisfies the \dagger first countability axiom (hence they are not metrizable), while the others are all metrizable. In (4) and (5), $\Delta_r = \Gamma$. Fig. 1 shows the relationship among the seven examples. Here $A \rightarrow B$ means that A lies over B , and $A \neq B$ means that in general neither $A \rightarrow B$ nor $B \rightarrow A$.

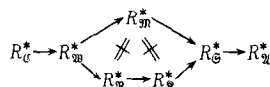


Fig. 1

D. Remarks

In contrast to topological compactifications (1)–(3), (4)–(7) can be regarded as potential-theoretic and have enough ideal boundary points so that one can solve the Dirichlet problem and introduce various measures there. Utilizing Green's functions, R. S. Martin [6] deduced the first important compactification and gave the integral representation of positive harmonic functions (the extension of the \dagger Poisson integral). Z. Kuramochi [3] obtained his compactification similarly by using N -Green functions introduced by himself instead of the usual Green's functions. In this compactification, the ideal boundary points can be considered to be interior points of the surfaces in a potential-theoretic sense. In case of finitely connected domains with smooth boundaries, both the Martin and Kuramochi boundaries coincide with the usual boundaries.

The Royden and Wiener compactifications were introduced as the maximal ideal spaces of respective function algebras. Their ideal boundaries contain extremely many points. However, these compactifications have elegant

properties and many applications. An analytic mapping φ of a Riemann surface R into another R' is called a Dirichlet (Fatou) mapping if φ can be extended to a continuous mapping of $R_{\mathfrak{A}}^*(R_{\mathfrak{B}}^*)$ to $R_{\mathfrak{A}}^*(R_{\mathfrak{B}}^*)$. For example, a Lindelöfian mapping (*AD*-function) is a Fatou (Dirichlet) mapping. The Dirichlet and Fatou mappings were investigated by Constantinescu, Cornea, and others.

By using the Martin boundary, Z. Kuramochi and M. Nakai proved the extension of the Evans-Selberg theorem to parabolic Riemann surfaces [3, 8]. The normal derivatives of *HD*-functions on the ideal boundaries and their applications were studied by Constantinescu, Cornea, F. Maeda, Y. Kusunoki, and others. The theory of compactification can be generalized to domains in \mathbf{R}^n , \dagger Green spaces, and harmonic spaces.

References

- [1] C. Constantinescu and A. Cornea, *Ideale Ränder Riemannscher Flächen*, Springer, 1963.
- [2] K. Hayashi, Sur une frontière des surfaces de Riemann, *Proc. Japan Acad.*, 37 (1961), 469–472.
- [3] Z. Kuramochi, Mass distributions on the ideal boundaries of abstract Riemann surfaces I, II, *Osaka Math. J.*, 8 (1956), 119–137, 145–186.
- [4] Y. Kusunoki, On a compactification of Green spaces, Dirichlet problem and theorems of Riesz type, *J. Math. Kyoto Univ.*, 1 (1962), 385–402.
- [5] Y. Kusunoki and S. Mori, On the harmonic boundary of an open Riemann surface I, *Japan. J. Math.*, 29 (1959), 52–56.
- [6] R. S. Martin, Minimal positive harmonic functions, *Trans. Amer. Math. Soc.*, 49 (1941), 137–172.
- [7] S. Mori, On a compactification of an open Riemann surface and its application, *J. Math. Kyoto Univ.*, 1 (1961), 21–42.
- [8] M. Nakai, On Evans potential, *Proc. Japan Acad.*, 38 (1962), 624–629.
- [9] M. Nakai, Radon-Nikodym densities between harmonic measures on the ideal boundary of an open Riemann surface, *Nagoya Math. J.*, 27 (1966), 71–76.
- [10] H. L. Royden, On the ideal boundary of a Riemann surface, *Contributions to the Theory of Riemann Surfaces*, Princeton Univ. Press, 1953, 107–109.
- [11] F. Maeda, M. Ohtsuka, et al., Kuramochi boundaries of Riemann surfaces, *Lecture notes in math.* 58, Springer, 1968.
- [12] L. Sario and M. Nakai, *Classification theory of Riemann surfaces*, Springer, 1970.

208 (X.7) Implicit Functions

A. General Remarks

Historically, a function y of x was called an implicit function of x if there was given a functional relation $f(x, y) = 0$ between x and y , but no explicit representation of y in terms of x (\rightarrow 165 Functions). Nowadays, however, the notion of implicit function is rigorously defined as follows: Suppose that a function $f(x_1, \dots, x_n, y)$ is of \dagger class C^1 in a domain G in the real $(n+1)$ -dimensional Euclidean space \mathbf{R}^{n+1} and that $f(x_1^0, \dots, x_n^0, y^0) = 0$, $f_y(x_1^0, \dots, x_n^0, y^0) \neq 0$ at a point $(x_1^0, \dots, x_n^0, y^0)$ in G . Then there is a unique function $g(x_1, \dots, x_n)$ of class C^1 in a neighborhood of the point (x_1^0, \dots, x_n^0) that satisfies $f(x_1, \dots, x_n, g(x_1, \dots, x_n)) = 0$, $y^0 = g(x_1^0, \dots, x_n^0)$ (**implicit function theorem**). The function g is called the **implicit function** determined by $f = 0$. The partial derivatives of g are given by the relation

$$\partial g / \partial x_j = -(\partial f / \partial x_j) / (\partial f / \partial y),$$

where $y = g(x_1, \dots, x_n)$. If the function f is of class C^r ($1 \leq r \leq \infty$ or $r = \omega$), then the function g is also of class C^r . In particular, when $n = 1$, letting x_1 be x , we have $dg/dx = -f_x/f_y$.

B. Jacobian Matrices and Jacobian Determinants

A mapping u from a domain G in \mathbf{R}^n into \mathbf{R}^m

$$u(x) = (u_1(x_1, \dots, x_n), \dots, u_m(x_1, \dots, x_n)),$$

$$x = (x_1, \dots, x_n),$$

is called a **mapping of class C^r** if each component u_1, \dots, u_m is of class C^r ($0 \leq r \leq \infty$ or $r = \omega$) in G . Given a mapping u of class C^1 from G into \mathbf{R}^m , we consider the following matrix, which gives rise to the differential du_x of the mapping u (\rightarrow 105 Differentiable Manifolds I):

$$\partial(u)/\partial(x) = (\partial u_j / \partial x_k)_{1 \leq j \leq m, 1 \leq k \leq n}. \quad (1)$$

This matrix is called the **Jacobian matrix** of the mapping u at x . If there is another mapping v of class C^1 from a domain containing the \dagger range U of u into \mathbf{R}^l , then we have the law of composition:

$$(\partial(v)/\partial(u))(\partial(u)/\partial(x)) = \partial(v)/\partial(x).$$

When $n = m$, the \dagger determinant of the matrix (1) is called the **Jacobian determinant** (or simply **Jacobian**), and is denoted by $D(u)/D(x)$, $D(u_1, \dots, u_n)/D(x_1, \dots, x_n)$ or

$$\frac{D(u_1, \dots, u_n)}{D(x_1, \dots, x_n)}.$$

Sometimes the notation ∂ is used instead of D ; but in the present article we distinguish the matrix from the determinant, using ∂ for the matrix and D for the determinant.

If $m=n$ and $D(u)/D(x)$ never vanishes at any point of the domain G , then u is called a **regular** (or **nonsingular**) **mapping of class C^1** . If the Jacobian $D(u)/D(x)$ is 0 at x , we say that u is **singular** at x . A mapping that is singular at every point in a set $S \subset G$ is said to be **degenerate** on S . For a regular mapping u , the sign of the Jacobian is constant in a connected domain G . If it is positive, the mapping u preserves the orientation of the coordinate system at each point in G , while if it is negative, the mapping changes the orientation. A point where u is degenerate is called a **critical point** of the mapping u , and its image under u is called a **critical value**. In general, the image of the mapping is "folded" along the set of critical points. The set of critical values of a mapping u of class C^1 (sending a domain in \mathbf{R}^n into \mathbf{R}^m) is of Lebesgue measure 0 in \mathbf{R}^m (**Sard's theorem**). If u is a regular mapping, then each point in the domain G of u has a neighborhood V such that the restriction of u on V is a \dagger topological mapping. Its inverse mapping $x(u)$ is also a regular mapping of class C^1 and satisfies the relation

$$\frac{D(u) D(x)}{D(x) D(u)} = 1$$

(**inverse mapping theorem**). If u is of class C^r ($1 \leq r \leq \infty$ or $r=\omega$), then so is its inverse mapping.

C. Functional Relations

A function $F(u_1, \dots, u_n)$ defined on a domain B in \mathbf{R}^n is called a **function with scattered zeros** if F has a zero point (i.e., there exists a point u for which $F(u)=0$) and if every open subset of B contains a point u such that $F(u) \neq 0$. Every \dagger analytic function $\neq 0$ has scattered zeros. Let $u(x)$ be a mapping from a domain G in \mathbf{R}^n into $B \subset \mathbf{R}^n$. Suppose that there exists a function $F(u)$ defined in B , of class C^r , with scattered zeros. If $F(u(x))=0$ for every x in G , then we say that the components u_1, \dots, u_n of the mapping u have a **functional relation of class C^r** or are **functionally dependent of class C^r** . In such a case, we sometimes say simply that u_1, \dots, u_n are **functionally dependent** or that they have a **functional relation**. If the components u_1, \dots, u_n of a mapping u of class C^1 are functionally dependent of class C^0 , then the Jacobian $D(u)/D(x)$ of u must vanish. Conversely, if the Jacobian $D(u)/D(x)$ of a mapping u of class C^1 is identically 0 in the domain G ,

the components u_1, \dots, u_n are functionally dependent of class C^∞ on every compact set in G (**Knopp-Schmidt theorem**) [1].

D. Implicit Functions Determined by Systems of Functions

Suppose that the \dagger rank of the Jacobian matrix of (1) is $r < m$ everywhere in G . Suppose that $u(x)$ is a mapping of class C^1 from a domain G in \mathbf{R}^n into \mathbf{R}^m , the Jacobian determinant $D(u_1, \dots, u_r)/D(x_1, \dots, x_r)$ never vanishes in G , and further $D(u_1, \dots, u_r, u_\rho)/D(x_1, \dots, x_r, x_\rho)$ is identically 0 in G for each ρ, σ with $r < \rho \leq m$, $r < \sigma \leq n$. Then the values $u_{r+1}(x), \dots, u_m(x)$ are determined by the values $u_1(x), \dots, u_r(x)$, and each u_ρ is represented as a function of class C^1 of u_1, \dots, u_r .

Let $u(x)$ be a mapping of class C^1 from a domain G in \mathbf{R}^n into \mathbf{R}^m and V the \dagger inverse image of a point u^0 . To study the properties of the set V , we assume, for simplicity, that u^0 is the origin. Suppose that the rank of the matrix $\partial(u)/\partial(x)$ is r for every point x in G and each of u_{r+1}, \dots, u_m is functionally dependent on u_1, \dots, u_r . Then each u_ρ ($r < \rho \leq m$) is a function $u_\rho(u_1, \dots, u_r)$ of u_1, \dots, u_r . The set V is empty if there is a ρ such that $u_\rho(0, \dots, 0) \neq 0$. On the other hand, if $u_\rho(0, \dots, 0)=0$ for all ρ ($r < \rho \leq m$), then V is the set of common zero points of the functions $u_1(x), \dots, u_r(x)$. Therefore, to study the set V , we can assume that $r=m \leq n$. If $m=n$, V consists of isolated points only. If $m < n$, then, changing the order of the variables x_1, \dots, x_n if necessary, we can assume that $D(u_1, \dots, u_m)/D(x_1, \dots, x_m) \neq 0$ at a point (x^0) in V . In this case, there is a unique function $\xi_\mu(x_{m+1}, \dots, x_n)$ of class C^1 ($1 \leq \mu \leq m$) in a neighborhood of (x_i^0) satisfying the following two conditions: (i) $x_\mu^0 = \xi_\mu(x_{m+1}^0, \dots, x_n^0)$; (ii) if the point (x_{m+1}, \dots, x_n) is in a neighborhood of $(x_{m+1}^0, \dots, x_n^0)$, then the point

$$(\xi_1(x_{m+1}, \dots, x_n), \dots, \xi_m(x_{m+1}, \dots, x_n),$$

$$x_{m+1}, \dots, x_n) \in V.$$

Each function ξ_μ is called an **implicit function** of x_{m+1}, \dots, x_n determined by the relations $u_1 = \dots = u_m = 0$. The \dagger total derivatives of the ξ_μ are determined from the system of linear equations

$$\sum_{k=1}^m \frac{\partial u_j}{\partial x_k} d\xi_k + \sum_{l=m+1}^n \frac{\partial u_j}{\partial x_l} dx_l = 0, \quad j=1, \dots, m.$$

The foregoing implicit function theorem is a local one. Among the **global implicit function theorems**, the following one, due to Hadamard, is useful: Let $x \mapsto y(x)$ be a mapping of class C^1 from \mathbf{R}^n into \mathbf{R}^n such that the inverse $\Phi^{-1}(x)$

of its Jacobian matrix $\Phi(x)$ is bounded in \mathbf{R}^n ; then the mapping is a \dagger diffeomorphism of class C^1 from \mathbf{R}^n onto \mathbf{R}^n .

The implicit function theorem holds also in complex spaces. Let $f_i(x_1, \dots, x_n, y_1, \dots, y_p)$, $1 \leq i \leq p$, be a system of \dagger holomorphic functions in a domain in the complex $(n+p)$ -dimensional space C^{n+p} . If (i) $f_i(x_1^0, \dots, x_n^0, y_1^0, \dots, y_p^0) = 0$, $1 \leq i \leq p$, and (ii) $D(f_1, \dots, f_p)/D(y_1, \dots, y_p)_{(x,y)=(x^0,y^0)} \neq 0$, then there exists a unique holomorphic solution $y_i = y_i(x_1, \dots, x_n)$ ($1 \leq i \leq p$) in a neighborhood of the point $x^0 = (x_1^0, \dots, x_n^0)$ that satisfies $f_i(x_1, \dots, x_n, y_1(x_1, \dots, x_n), \dots, y_p(x_1, \dots, x_n)) = 0$ ($1 \leq i \leq p$).

E. Linear Relations

Suppose that $u(x)$ is a mapping of class C^{m-1} from \mathbf{R}^1 into \mathbf{R}^m . Its components (u_1, \dots, u_m) are functions of class C^{m-1} . Then the determinant

$$\begin{vmatrix} u_1 & u_2 & \dots & u_m \\ u'_1 & u'_2 & \dots & u'_m \\ \vdots & \vdots & \ddots & \vdots \\ u_1^{(m-1)} & u_2^{(m-1)} & \dots & u_m^{(m-1)} \end{vmatrix}$$

is called the **Wronskian determinant** (or simply the **Wronskian**) of the functions u_1, \dots, u_m and is denoted by $W(u_1, u_2, \dots, u_m)$. If the functions u_1, \dots, u_m are \dagger linearly dependent, i.e., if there exist constants c_j not all zero satisfying $\sum_{j=1}^m c_j u_j(x) = 0$ identically, then the Wronskian vanishes identically. Therefore, if $W(u_1, \dots, u_m) \neq 0$, then the functions u_1, \dots, u_m are linearly independent. Conversely, if $W(u_1, \dots, u_m) = 0$ identically, and further if there is at least one nonvanishing Wronskian for $u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_m$ ($1 \leq i \leq m$), then the functions u_1, \dots, u_m are linearly dependent. The necessity of the additional condition is shown by the following example: $u_1 = x^3$ and $u_2 = |x|^3$ are of class C^1 in the interval $[-1, 1]$ and linearly independent, but they satisfy $W(x^3, |x|^3) = 0$ identically. However, the additional condition is unnecessary if the functions u_1, \dots, u_m are analytic. Similar theorems are valid in a domain in a complex plane.

Furthermore, if $u(x)$ is a continuous mapping from an interval $[a, b]$ in \mathbf{R}^1 into \mathbf{R}^m , the determinant

$$G(u_1, \dots, u_m) = \begin{vmatrix} (1, 1) & \dots & (1, m) \\ (2, 1) & \dots & (2, m) \\ \vdots & \ddots & \vdots \\ (m, 1) & \dots & (m, m) \end{vmatrix},$$

$$(j, k) = \int_a^b u_j(x) u_k(x) dx$$

is called the **Gramian determinant** (or simply **Gramian**). The Gramian is the \dagger discriminant of

the quadratic form

$$\int_a^b \left(\sum_{j=1}^m \xi_j u_j(x) \right)^2 dx$$

of ξ_1, \dots, ξ_m and is equal to

$$\frac{1}{m!} \int_a^b \dots \int_a^b (\det(u_j(x_k)))^2 dx_1 \dots dx_m.$$

We always have $G(u_1, \dots, u_m) \geq 0$, and $G(u_1, \dots, u_m) = 0$ if and only if u_1, \dots, u_m are linearly dependent. The Gramian is defined if the functions u_1, \dots, u_m are \dagger square integrable in the sense of Lebesgue. In that case, the condition $G(u_1, \dots, u_m) = 0$ holds if and only if u_1, \dots, u_m are linearly dependent \dagger almost everywhere, i.e., there are constants c_1, \dots, c_m not all zero such that the relation $c_1 u_1(x) + \dots + c_m u_m(x) = 0$ holds except on a set of Lebesgue measure 0.

References

- [1] K. Knopp and R. Schmidt, Funktional-determinanten und Abhängigkeit von Funktionen, Math. Z., 25 (1926), 373–381.
- [2] G. Doetsch, Die Funktionaldeterminante als Deformationsmass einer Abbildung und als Kriterium der Abhängigkeit von Funktionen, Math. Ann., 99 (1928), 590–601.
- [3] W. Rudin, Principles of mathematical analysis, McGraw-Hill, second edition, 1964.
- [4] H. Cartan, Calcul différentiel, Hermann, 1967.
- [5] J. T. Schwartz, Nonlinear functional analysis, Gordon & Breach, 1969.

209 (XXI.5) Indian Mathematics

India was one of the earliest civilizations, but because it has no precise chronological record of ancient times, it is said to possess no history. Indian mathematics seems to have developed under the influence of the cult of Brahma, as did the calendar. It may also have some relation to the mathematics of the Near East and China, but this is difficult to trace. The word *ganita* (computation) appears in early religious writings; after the beginning of the Christian Era, it was classified into *pātī-gaṇita* (arithmetic), *bīja-gaṇita* (algebra), and *kṣetra-gaṇita* (geometry), thus showing some systematization. The Buddhists (notably Nagarjuna) had a kind of logic, but it had no re-

lation to mathematics. Unlike the Greeks, the Indians had no demonstrational geometry, but they had symbolic algebra and a position system of numeration.

Indian geometry was computational; Āryabhaṭa (c. 476–c. 550) computed the value of π as 3.1416; Brahmagupta (598–c. 660) had a formula to compute the area of quadrangles inscribed in a circle; and Bhāskara (1114–1185) gave a proof of the Pythagorean theorem. In trigonometry, Āryabhaṭa made a table of sines of angles between 0° and 90° for every 3.75° interval. The name “sine” is related to the Sanskrit *jya*, which referred to half of the chord of the double arc.

The Indians had a remarkable system of algebra. At the beginning they had no operational symbols and described in words the rules for solving equations. Brahmagupta worked on the \dagger Pell equation $ax^2 + 1 = y^2$. Bhāskara knew that a quadratic equation can have two roots that can be positive and negative, but did not assign any meaning to the negative root in such cases. Bhāskara also introduced algebraic symbols.

The symbol 0 was used in India from about 200 B.C. to denote the void place in the position system of numeration; 0 as a number is found in a book by Bakhshālī published in the 3rd century A.D. The number 0 is defined as $a - a = 0$ in our notation, and the rules $a \pm 0 = a$, $0 \times a = 0$, $\sqrt{0} = 0$, $0 \div a = 0$ are mentioned. Brahmagupta prohibited division by 0 in arithmetic, but in algebra he called the “quantity” $a \div 0$ *taccheda*. Bhāskara called it *kahara* and made it play a role similar to that of our infinity.

Some historians assert that the Indians had the ideas of infinity and infinitesimal. Some hold that the Indian position system of numeration arose from the circumstance that the names of numbers differed according to their positions. The Indian numeration system was exported to Europe through Arabia, and had great influence on the development of mathematics.

References

- [1] M. Cantor, *Vorlesungen über Geschichte der Mathematik I*, Teubner, third edition, 1907.
- [2] G. Sarton, *Introduction to the history of science I, From Homer to Omar Khayyam* Carnegie Institute of Washington, 1927.
- [3] B. Datta and A. N. Singh, *History of Hindu mathematics I, II*, Lahore, 1935–1938. (Asia Publishing House, 1962).

210 (II.25) Inductive Limits and Projective Limits

A. General Remarks

Inductive and projective limits can be defined over any \dagger preordered set I and in any \dagger category. We first explain the definition of these limits in the special case where I is a \dagger directed set and the category is that of sets, of groups, or of topological spaces. The simplest case is when I is the ordered set \mathbb{N} of the natural numbers.

B. The Limit of Sets

Let I be a directed set. Suppose that we are given a set X_i for each $i \in I$ and a mapping $\varphi_{ji}: X_i \rightarrow X_j$ for each pair (i, j) of elements of I with $i \leq j$, such that $\varphi_{ii} = 1_{X_i}$ (the identity mapping on X_i) and $\varphi_{ki} = \varphi_{kj} \circ \varphi_{ji}$ ($i \leq j \leq k$). Then we denote the system by (X_i, φ_{ji}) and call it an **inductive system** (or **direct system**) of sets over I . Let S be the \dagger direct sum $\coprod_i X_i$ of the sets X_i ($i \in I$), and define an equivalence relation in S as follows: $x \in X_i$ and $y \in X_j$ are equivalent if and only if there exists a $k \in I$ such that $i \leq k$, $j \leq k$, $\varphi_{ki}(x) = \varphi_{kj}(y)$. Let D be the \dagger quotient set of S by this equivalence relation, and let $f_i: X_i \rightarrow D$ ($i \in I$) be the canonical mappings. Then we have I(1) $f_j \circ \varphi_{ji} = f_i$ ($i \leq j$); I(2) for any set X , and for any system of mappings $g_i: X_i \rightarrow X$ ($i \in I$) satisfying $g_j \circ \varphi_{ji} = g_i$ ($i \leq j$), there exists a unique mapping $f: D \rightarrow X$ such that $f \circ f_i = g_i$ ($i \in I$). We call (D, f_i) the **inductive limit** (or **direct limit**) of the inductive system (X_i, φ_{ji}) over I , and denote it by $\lim X_i$ or $\text{ind } \lim X_i$ (more precisely, by $\varinjlim_{i \in I} X_i$ or $\text{ind } \lim_{i \in I} X_i$).

Suppose, dually, that we are given a set X_i for each $i \in I$ and a mapping $\psi_{ij}: X_j \rightarrow X_i$ for each $i \leq j$, such that $\psi_{ii} = 1_{X_i}$ and $\psi_{ik} = \psi_{ij} \circ \psi_{jk}$ ($i \leq j \leq k$). Then we denote the system by (X_i, ψ_{ij}) and call it a **projective system** (or **inverse system**) of sets over I . Let P be the subset of the Cartesian product $\prod X_i$ defined by $P = \{(x_i) \mid \psi_{ij}(x_j) = x_i \text{ (} i \leq j \text{)}\}$, and let $p_i: P \rightarrow X_i$ be the canonical mappings. Then we have P(1) $\psi_{ij} \circ p_j = p_i$ ($i \leq j$); P(2) for any set X , and for any system of mappings $q_i: X \rightarrow X_i$ satisfying $\psi_{ij} \circ q_j = q_i$ ($i \leq j$), there exists a unique mapping $p: X \rightarrow P$ such that $p_i \circ p = q_i$ ($i \in I$). We call (P, p_i) the **projective limit** (or **inverse limit**) of the projective system (X_i, ψ_{ij}) over I and denote it by $\lim X_i$ or $\text{proj } \lim X_i$.

Note that we may replace I by any cofinal subset of I without changing the limits.

C. The Limit of Groups and of Topological Spaces

If, in the notation of Section B, X_i is a group and $\varphi_{ji}(\psi_{ij})$ is a homomorphism, then we say that (X_i, φ_{ji}) ((X_i, ψ_{ij})) is an **inductive (projective) system of groups**. The inductive limit (as a set) $D = \varinjlim X_i$ has the structure of a group for which the canonical mappings f_i are homomorphisms. With this group structure, D is called the **inductive limit (group)** of the inductive system of groups. It satisfies properties I(1) and I(2) with group X and homomorphisms g_i and f . Similarly, the projective limit (as a set) $P = \varprojlim X_i$ has a unique group structure such that each $p_i: P \rightarrow X_i$ is a homomorphism, namely, that of a subgroup of the direct product group $\prod X_i$. The group P is called the **projective limit (group)** of the projective system of groups. When each X_i is a module over a fixed ring A , we get entirely similar results by considering A -homomorphisms instead of group homomorphisms.

Next, let X_i be a topological space and φ_{ji} and ψ_{ij} be continuous mappings. Then (X_i, φ_{ji}) ((X_i, ψ_{ij})) is called an **inductive (projective) system of topological spaces**. If we introduce in $D = \varinjlim X_i$ the topology of a quotient space of the topological direct sum of the spaces X_i ($i \in I$), then the f_i are continuous, and I(1) and I(2) hold with sets and mappings replaced by topological spaces and continuous mappings. Similarly, if we view $P = \varprojlim X_i$ as a subspace of the \dagger product space $\prod X_i$, then the p_i are continuous and P(1) and P(2) hold with the same modification as before. The spaces D and P are called the **inductive limit (space)** and the **projective limit (space)** of the system of topological spaces, respectively. The projective limit of Hausdorff (compact) spaces is also Hausdorff (compact).

Furthermore, if the X_i ($i \in I$) form a topological group and φ_{ji}, ψ_{ij} are continuous homomorphisms, then $\varinjlim X_i$ and $\varprojlim X_i$ are topological groups, and properties I(1), I(2), P(1), and P(2) are satisfied for topological groups and continuous homomorphisms (\rightarrow 423 Topological Groups). In particular, projective limits of finite groups are \dagger totally disconnected compact groups and are called **profinite groups**; they occur, e.g., as the ring of $\dagger p$ -adic integers and as the \dagger Galois group of an infinite \dagger Galois extension. Conversely, the \dagger germs of continuous functions at a point x in a topological space X , and other kinds of germs (\rightarrow 383 Sheaves), are important examples of inductive limits of groups.

D. Limits in a Category

Let I be a preordered set and \mathcal{C} a category. If we are given an object X_i of a category \mathcal{C} for each $i \in I$ and a \dagger morphism $\varphi_{ji}: X_i \rightarrow X_j$ of \mathcal{C} for each pair (i, j) of elements of I with $i \leq j$, and if the conditions $\varphi_{ii} = 1_{X_i}$, $\varphi_{ki} = \varphi_{kj} \circ \varphi_{ji}$ ($i \leq j \leq k$) are satisfied, then we call the system (X_i, φ_{ji}) an **inductive system** over I in the category \mathcal{C} . A **projective system** in \mathcal{C} is defined dually: It is an inductive system in the \dagger dual category \mathcal{C}° . If we view I as a category (\rightarrow 52 Categories and Functors B), then an inductive (projective) system in \mathcal{C} over the index set I is a \dagger covariant (\dagger contravariant) functor from I to \mathcal{C} . Now if an object $D \in \mathcal{C}$ and morphisms $f_i: X_i \rightarrow D$ ($i \in I$) satisfy conditions I(1) and I(2) with the modification that X is an object and g_i, f are morphisms in \mathcal{C} , then the system (D, f_i) is called the **inductive limit** of (X_i, φ_{ji}) and is denoted by $\varinjlim X_i$. Similarly, if an object $P \in \mathcal{C}$ and morphisms $p_i: P \rightarrow X_i$ ($i \in I$) satisfy P(1) and P(2) with a similar modification, then (P, p_i) is called the **projective limit** of (X_i, ψ_{ij}) and is written $\varprojlim X_i$. By I(2) and P(2), these limits are unique if they exist.

In the categories of sets, of groups, of modules, and of topological spaces, inductive and projective limits always exist. Note that if the ordering of I is such that $i \leq j$ implies $i = j$, i.e., if there is no ordering between two distinct elements of I , then the inductive (projective) limit is the \dagger direct sum (\dagger direct product) (\rightarrow 52 Categories and Functors E).

Let (X_i, φ_{ji}) , (X'_i, φ'_{ji}) be two inductive systems over the same index set I , and let $\varphi_i: X_i \rightarrow X'_i$ ($i \in I$) be morphisms satisfying $\varphi'_{ji} \circ \varphi_i = \varphi_i \circ \varphi_{ji}$ ($i \leq j$). Then the system (φ_i) is called a **morphism** between the inductive systems. Such a morphism is a \dagger natural transformation between the inductive systems viewed as functors $I \rightarrow \mathcal{C}$. If $\varinjlim X_i$ and $\varinjlim X'_i$ exist, then (φ_i) induces a morphism $\varinjlim \varphi_i: \varinjlim X_i \rightarrow \varinjlim X'_i$ in a natural way, and similarly for projective limits.

For the more abstract notion of limit of a functor \rightarrow [5]. For the theory of procategories \rightarrow [4].

References

- [1] S. Lefschetz, Algebraic topology, Amer. Math. Soc. Colloq. Publ., 1942.
- [2] S. Eilenberg and N. Steenrod, Foundations of algebraic topology, Princeton Univ. Press, 1952.
- [3] H. Cartan and S. Eilenberg, Homological algebra, Princeton Univ. Press, 1956.
- [4] A. Grothendieck, Technique de descente et théorèmes d'existence en géométrie algébrique II. Le théorème d'existence en théorie formelle

des modules, Sém. Bourbaki, Exposé 195, 1959/1960 (Benjamin, 1966).
[5] M. Artin, Grothendieck topology, Lecture notes, Harvard Univ. 1962.

211 (X.3) Inequalities

A. General Remarks

In this article we consider various properties of inequalities between real numbers. An inequality that holds for every real number (e.g., $x^2 \geq 0$) is called an **absolute inequality**; otherwise it is called a **conditional inequality**. When we are given a conditional inequality (e.g., $x(x-1) < 0$), the set of all real numbers that satisfy it is called the **solution** of the inequality. The process of obtaining a solution is called **solving** the conditional inequality.

B. Solution of a Conditional Inequality

Suppose that a conditional inequality is given by $f(x) > 0$ (or $f(x) \geq 0$), where f is a continuous function defined for every real number. If the equation $f(x) = 0$ has no solution, then we have either $f(x) > 0$ or $f(x) < 0$ for all x . On the other hand, if α and β ($\alpha < \beta$) are adjacent roots of the equation $f(x) = 0$, the sign of $f(x)$ is unchanged in the open interval (α, β) . Therefore the solution of the given inequality depends essentially on the solution of the equation $f(x) = 0$. If inequalities involve two variables x, y and are given by $f(x, y) > 0$, $g(x, y) > 0$ for continuous functions f and g , the solution is, in general, a domain in the xy -plane bounded by the curves $f(x, y) = 0$ and $g(x, y) = 0$. Similar results hold for the case of inequalities involving more than two variables.

C. Famous Absolute Inequalities

(1) Inequalities concerning **means** (or **averages**): Suppose that we are given an n -tuple $a = (a_1, \dots, a_n)$, $a_r \geq 0$. We set

$$M_r = M_r(a) = \left(\frac{1}{n} \sum_{v=1}^n a_v^r \right)^{1/r}.$$

If at least one a_r is 0 and $r < 0$, we put $M_r = 0$. In particular, we put

$$A = M_1,$$

$$G = \lim_{r \rightarrow \infty} M_r = \left(\prod_{v=1}^n a_v \right)^{1/n},$$

$$H = M_{-1};$$

these are called the **arithmetic mean**, **geometric mean**, and **harmonic mean** of a_v ($v = 1, \dots, n$), respectively. Except when either all a_v are identical or some a_v is 0 and $r \leq 0$, the function M_r increases [†]strictly monotonically as r increases, and $M_r \rightarrow \min a_v$ ($r \rightarrow -\infty$), $M_r \rightarrow \max a_v$ ($r \rightarrow +\infty$). Therefore we always have $\min a_v \leq M_r \leq \max a_v$. In particular, we have $H < G < A$ if the a_v are all positive and not all equal.

Let $p(x) (> 0)$, $f(x) (\geq 0)$ be [†]integrable functions on a [†]measurable set E . We put

$$M_r(f) = \left(\int_E p f^r dx / \int_E p dx \right)^{1/r}, \quad r \neq 0.$$

Furthermore, if $M_r(f)$ is strictly positive for some $r > 0$, we put

$$\begin{aligned} M_0(f) &= \lim_{r \rightarrow +0} M_r(f) \\ &= \exp \left(\int_E p \log f dx / \int_E p dx \right). \end{aligned}$$

We call $M_r(f)$ the **mean of degree r** of the function $f(x)$ with respect to the **weight function** $p(x)$. It has properties similar to those of $M_r(a)$. In particular, when the weight function $p = 1$, the means $M_1(f)$, $M_0(f)$, $M_{-1}(f)$ are called the **arithmetic mean**, **geometric mean**, and **harmonic mean** of f , respectively.

(2) The Hölder inequality: Suppose that $p \neq 0, 1$ and $(p-1)(q-1) = 1$; that is, $1/p + 1/q = 1$, and $a_v > 0$, $b_v > 0$. Then, in general, we have the **Hölder inequality**:

$$\sum_v a_v b_v \leq \left(\sum_v a_v^p \right)^{1/p} \left(\sum_v b_v^q \right)^{1/q}, \quad p \leq 1,$$

where the inequality signs in the first inequality are taken in accordance with $p < 1$ or $p > 1$. The summation may be infinite if the sums are convergent. The inequality sign is replaced by the equality sign if and only if there exist constant factors λ and μ such that $\lambda a_v^p = \mu b_v^q$ for all v . The Hölder inequality for $p = q = 2$ is called the **Cauchy inequality** (or **Cauchy-Schwarz inequality**).

For two measurable positive functions $f(x)$, $g(x)$, we have the **Hölder integral inequality**:

$$\int_E f g dx \leq \left(\int_E f^p dx \right)^{1/p} \left(\int_E g^q dx \right)^{1/q}, \quad p \leq 1,$$

except when there exist two constant factors λ and μ ($(\lambda, \mu) \neq (0, 0)$) such that $\lambda f^p = \mu g^q$ holds [†]almost everywhere. The above inequality is replaced by equality if and only if we are in this exceptional case. The case where $p = q = 2$ is called the **Schwarz inequality** (or **Bunyakovskii inequality**).

(3) The Minkowski inequality: Suppose that $p \neq 0, 1$ and $a_v > 0$, $b_v > 0$. Then we have the

Minkowski inequality:

$$\left(\sum_v (a_v + b_v)^p\right)^{1/p} \geq \left(\sum_v a_v^p\right)^{1/p} + \left(\sum_v b_v^p\right)^{1/p},$$

$$p \leq 1,$$

except when $\{a_v\}$ and $\{b_v\}$ are proportional. The inequality is replaced by equality if and only if we are in the exceptional case.

The corresponding integral inequality for positive functions $f(x)$, $g(x)$ is

$$\left(\int_E (f+g)^p dx\right)^{1/p} \geq \left(\int_E f^p dx\right)^{1/p} + \left(\int_E g^p dx\right)^{1/p}, \quad p \leq 1,$$

except when $f(x)/g(x)$ is constant almost everywhere. The inequality is replaced by equality if and only if we are in the exceptional case.

D. Related Topics

Absolute inequalities are important in analysis, especially in connection with techniques to prove convergence or for error estimates. However, there seldom are general principles for deriving such inequalities, except for a few elementary theorems.

For other famous inequalities → Appendix A, Table 8. For related topics → 88 Convex Analysis; for convex functions and their applications → 255 Linear Programming; for linear inequalities → 212 Inequalities in Physics.

References

- [1] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge Univ. Press, first edition, 1934; revised edition, 1952.
- [2] V. I. Levin and S. B. Stechkin (Stechkin), *Inequalities*, Amer. Math. Soc. Transl., (2) 14 (1960), 1–29 (English translation of the appendix to the Russian translation of [1]).
- [3] E. F. Beckenbach and R. Bellman, *Inequalities*, Erg. Math., Springer, second revised printing, 1965.
- [4] E. F. Beckenbach and R. Bellman, *An introduction to inequalities*, Random House, 1961.
- [5] N. D. Kazarinoff, *Geometric inequalities*, Random House, 1961.
- [6] N. D. Kazarinoff, *Analytic inequalities*, Holt, Rinehart and Winston, 1961.
- [7] W. Walter, *Differential- und Integral-Ungleichungen*, Springer, 1964.
- [8] O. Shisha (ed.), *Inequalities*, Academic Press, I, 1967; II, 1970; III, 1972.

**212 (XX.36)
Inequalities in Physics****A. Correlation Inequalities**

Let μ be a probability measure on a space X (with σ -field \mathcal{B}) and f_j be measurable functions on X , and write

$$\langle f_1 \dots f_n \rangle = \int_X f_1(x) \dots f_n(x) d\mu(x).$$

A number of inequalities among such expectations under a variety of conditions on the measure μ and functions f_j are known as correlation inequalities, after their original occurrence for correlation functions in the statistical mechanics of lattice gases.

The earliest results were obtained by Griffiths (*J. Math. Phys.*, 8 (1967)), for the case where X is the product $\prod_{i=1}^N I_i$ of two point sets $I_i = \{1, -1\}$, the f are k th coordinate functions $\sigma_k = \sigma_k(x)$ ($= \pm 1$) of $x \in X$, $d\mu(x)$ is the counting measure multiplied by a Gibbs factor $Z^{-1} e^{-\beta H(x)}$ with a ferromagnetic Hamiltonian

$$H(x) = - \sum_{i < j} J_{ij} \sigma_i(x) \sigma_j(x), \quad J_{ij} = J_{ji} \geq 0,$$

$\beta \geq 0$, and the normalization factor $Z = \sum_x e^{-\beta H(x)}$. His conclusions are

$$\langle \sigma_k \sigma_l \rangle \geq 0 \quad (\text{Griffiths's first inequality}),$$

$$\langle \sigma_k \sigma_l \sigma_m \sigma_n \rangle \geq \langle \sigma_k \sigma_l \rangle \langle \sigma_m \sigma_n \rangle$$

(Griffiths's second inequality).

These were extended by Kelly and Sherman (*J. Math. Phys.*, 9 (1968)) as

$$\langle \sigma_A \rangle \geq 0 \quad (\text{GKS first inequality}),$$

$$\langle \sigma_A \sigma_B \rangle \geq \langle \sigma_A \rangle \langle \sigma_B \rangle$$

(GKS second inequality),

where A and B are subsets of $\{1, \dots, N\}$, $\sigma_A = \prod_{i \in A} \sigma_i$, and $H = - \sum_{A \subset N} J_A \sigma_A$ with $J_A \geq 0$. A further generalization (for example, $I_i = \mathbf{R}$) can be found in [1, 2].

Under the same situation with

$$H = - \sum_{i < j} J_{ij} \sigma_i \sigma_j - \sum_{i=1}^N h_i \sigma_i, \quad J_{ij} \geq 0, \quad h_i \geq 0,$$

the following GHS inequality by Griffiths, Hurst, and Sherman (*J. Math. Phys.*, 11 (1970)) holds:

$$\begin{aligned} \langle \sigma_i \sigma_j \sigma_k \rangle - \langle \sigma_i \rangle \langle \sigma_j \sigma_k \rangle - \langle \sigma_j \rangle \langle \sigma_i \sigma_k \rangle \\ - \langle \sigma_k \rangle \langle \sigma_i \sigma_j \rangle + 2 \langle \sigma_i \rangle \langle \sigma_j \rangle \langle \sigma_k \rangle \leq 0. \end{aligned}$$

If $h_i = 0$ for all i , then

$$\begin{aligned} \langle \sigma_i \sigma_j \sigma_k \sigma_l \rangle - \langle \sigma_i \sigma_l \rangle \langle \sigma_j \sigma_k \rangle - \langle \sigma_j \sigma_l \rangle \langle \sigma_i \sigma_k \rangle \\ - \langle \sigma_k \sigma_l \rangle \langle \sigma_i \sigma_j \rangle + 2 \langle \sigma_i \sigma_l \rangle \langle \sigma_j \sigma_l \rangle \langle \sigma_k \sigma_l \rangle \leq 0. \end{aligned}$$

Further generalizations can be found in [2, 3], and the references quoted therein.

Let X be a finite distributive lattice, μ be a positive measure satisfying the condition $\mu(x \wedge y)\mu(x \vee y) \geq \mu(x)\mu(y)$, and f and g be both increasing or both decreasing functions on X . Then the following FKG inequality by Fortuin, Kasteleyn, and Ginibre (*Comm. Math. Phys.*, 22 (1971)) holds:

$$\langle fg \rangle \geq \langle f \rangle \langle g \rangle.$$

B. Inequalities Involving Traces of Matrices

Let $\text{tr } A$ denote the trace of a matrix A . Some of the earlier results applied to statistical mechanics are as follows, where $\rho \geq 0$, $\sigma \geq 0$, $A^* = A$, $B^* = B$:

$$\text{tr}(\rho \log \rho - \rho \log \sigma - \rho + \sigma) \geq 0$$

(Klein inequality),

$$\text{tr}(e^{A+B})/\text{tr } e^A \geq \exp(\text{tr } e^A B / \text{tr } e^A)$$

(Peierls-Bogolyubov inequality),

$$\text{tr}(e^A e^B) \geq \text{tr}(e^{A+B})$$

(Golden-Thompson inequality).

The following inequality is related to the last inequality and was proved for a general $m \geq 0$ by Lieb and Thirring (*Studies in Mathematical Physics*, Lieb, Simon, and Wightman (eds.), Princeton Univ. Press, 1976):

$$\text{tr}(\rho \sigma)^m \leq \text{tr}(\rho^m \sigma^m), \quad \rho \geq 0, \quad \sigma \geq 0.$$

For the Hilbert-Schmidt norm $\|A\|_{\text{H.S.}} \equiv (\text{tr } A^* A)^{1/2}$ and the trace-class norm $\|A\|_{\text{tr}} \equiv \text{tr}|A|$, where $|A| = (A^* A)^{1/2}$, the Powers-Størmer inequality (*Comm. Math. Phys.*, 16 (1970)) holds:

$$\|\rho^2 - \sigma^2\|_{\text{tr}} \geq (\|\rho - \sigma\|_{\text{H.S.}})^2, \quad \rho \geq 0, \quad \sigma \geq 0.$$

Araki and Yamagami (*Comm. Math. Phys.*, 81 (1981)) give

$$\| |C| - |D| \|_{\text{H.S.}} \leq \sqrt{2} \|C - D\|_{\text{H.S.}}$$

If $C^* = C$ and $D^* = D$, then $\sqrt{2}$ can be removed.

The entropy $S(\rho) = -\text{tr } \rho \log \rho$ is a concave function of $\rho \geq 0$ satisfying the triangular inequality (Araki and Lieb, *Comm. Math. Phys.*, 18 (1970)):

$$S(\rho_1) + S(\rho_2) \geq S(\rho_{12}) \geq |S(\rho_1) - S(\rho_2)|$$

and strong subadditivity (Lieb and Ruskai, *J. Math. Phys.*, 14 (1973)):

$$S(\rho_{123}) - S(\rho_{12}) - S(\rho_{13}) + S(\rho_1) \leq 0,$$

where ρ_{123} is a matrix on the tensor product space $H_1 \otimes H_2 \otimes H_3$, $\rho_{jk} = \text{tr}_i \rho_{123}$, $\rho_i = \text{tr}_j \text{tr}_k \rho_{123}$ ($\{i, j, k\} = \{1, 2, 3\}$) and tr_j is a (partial) trace taken only on the space H_j . The

last inequality is based on the concavity of the function $f(\rho) = \text{tr} \exp(A + \log \rho)$ in $\rho \geq 0$ ($A^* = A$) proved by Lieb [4] (also see Epstein, *Comm. Math. Phys.*, 31 (1973)), who also proved the joint concavity of $\text{tr}(C^* \rho^s C \sigma^r)$ ($r \geq 0, s \geq 0, r + s \leq 1$) in $\rho \geq 0$ and $\sigma \geq 0$. This concavity for the case $r = s = 1/2$ was previously proved by Wigner and Yanase (*Proc. Nat. Acad. Sci. US*, 49 (1963)), and the general case has been conjectured by Wigner, Yanase, and Dyson. It leads to the joint concavity of the relative entropy $S(\rho, \sigma) = \text{tr } \rho(\log \rho - \log \sigma)$ (defined to be $+\infty$ if $\sigma \psi = 0$ and $\rho \psi \neq 0$ for some vector ψ) in $\rho \geq 0$ and $\sigma \geq 0$ as well as its monotonicity $S(\alpha(\rho), \alpha(\sigma)) \leq S(\rho, \sigma)$ for any trace-preserving expectation mapping α . (For entropy, see also [5].)

The above results have generalizations in the context of von Neumann algebras (Araki, *Publ. Res. Inst. Math. Sci.*, 11 (1975); 13 (1977); Uhlmann, *Comm. Math. Phys.*, 54 (1977); [6]).

C. Operator Monotone and Operator Convex Functions

A real-valued function $f(x)$ defined on an interval I (finite or infinite; open, half-open, or closed) is called matrix monotone increasing (decreasing) of order m if $f(A) \geq f(B)$ whenever $m \times m$ Hermitian matrices A and B with their eigenvalues contained in I satisfy $A \geq B$ ($A \leq B$) and is called operator monotone if it is matrix monotone of order n for all positive integers n . $f(x)$ is called matrix convex of order m if $f[(1-t)A + tB] \leq (1-t)f(A) + tf(B)$ for $0 \leq t \leq 1$ and all $m \times m$ Hermitian matrices A and B with their eigenvalues in I and is called operator convex if it is matrix convex of order n for any positive integer n . The functions x^α ($0 \leq \alpha \leq 1$), $\log x$, and $-(x+a)^{-1}$ ($a \geq 0$) are all operator monotone increasing in the half-line interval $x > 0$. The functions $(x+a)^{-1}$ and $x \log x$ are operator convex in the same interval.

A function $f(x)$ is operator monotone increasing in an open interval (a, b) if and only if it is analytic in (a, b) and has an analytic continuation to the whole upper half-plane where the function has a nonnegative imaginary part [7]. Another necessary and sufficient condition is

$$\sum_{i,k=0}^N f^{(i+k+1)}(x) \zeta_i \zeta_k / (i+k+1)! \geq 0$$

for all $x \in (a, b)$, for all real ζ , and for all positive integers N . An operator monotone function $f(x)$ on an open interval $(-R, R)$ has the integral representation

$$f(x) = f(0) + f'(0) \int x/(1-tx) d\mu(t),$$

where μ is a probability measure with its support contained in $[-R^{-1}, R^{-1}]$ and, if f is not a constant, is uniquely determined by f . The set of all extremal points of the set of all operator monotone increasing functions on $(-1, 1)$ satisfying $f(0)=0$ and $f'(0)=1$ is exactly the set of functions $x/(1-tx)$, $|t| \leq 1$, which appear as integrands in the above formula. An operator monotone function on \mathbf{R} must be linear.

An operator convex function in an open interval $(-R, R)$ has the integral representation

$$f(x) = f(0) + f'(0)x + (f''(0)/2) \int x^2/(1-tx) d\mu(t),$$

where μ is a probability measure with its support contained in $[-R^{-1}, R^{-1}]$ and, if f is not linear, is uniquely determined by f . An operator convex function in \mathbf{R} must be at most a quadratic polynomial satisfying $f''(0) \geq 0$.

For a continuous real function on an interval $[0, \alpha)$ ($0 \leq \alpha < \infty$), the following four conditions are mutually equivalent: (i) f is operator convex and $f(0) \leq 0$, (ii) $f(a^*Aa) \leq a^*f(A)a$ for any self-adjoint operator A with its spectrum in $[0, \alpha)$ and any operator a with its norm not exceeding 1, (iii) the preceding condition with a limited to projections, (iv) $x^{-1}f(x)$ is operator monotone increasing in an open interval $(0, \alpha)$. (See [8–11] and Hansen and Pedersen, Jensen's inequality for operators and Löwner's theorem, *Math. Ann.*, 258 (1982), 229–241.

References

- [1] J. Ginibre, General formulation of Griffiths' inequality, *Comm. Math. Phys.*, 16 (1970), 310–328.
- [2] B. Simon, *Functional integration and quantum physics*, Academic Press, 1979, ch. 4.
- [3] C. Newman, Gaussian correlation inequalities, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 33 (1975), 75–93.
- [4] E. H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Advances in Math.*, 11 (1973), 267–288.
- [5] A. Wehrl, General properties of entropy, *Rev. Mod. Phys.*, 50 (1978), 221–260.
- [6] H. Araki, Inequalities in von Neumann algebras, CNRS Publication R.C.P., 25 vols., 22 (1975), 1–25.
- [7] K. Löwner, Über monotone Matrixfunktionen, *Math. Z.*, 38 (1934), 177–216.
- [8] J. Bendat and S. Sherman, Monotone and convex operator functions, *Trans. Amer. Math. Soc.*, 79 (1955), 58–71.
- [9] W. Donoghue, *Monotone matrix functions and analytic continuation*, Springer, 1974.
- [10] C. Davis, Notions generalizing convexity for functions defined on spaces of matrices, *Amer. Math. Soc. Proc. Symposia in Pure Math.*, 7 (1963), 187–201.
- [11] F. Kraus, Über konvexe Matrixfunktionen, *Math. Z.*, 41 (1936), 18–42.

213 (XIX.12) Information Theory

A. General Remarks

The mathematical theory of information transmission in communication systems, first developed by C. E. Shannon [1] and now called **information theory**, is one of the most important fields of mathematical science. It consists of two main themes, **channel coding theory** and **source coding theory**. The purpose of channel coding theory is to ascertain the existence of schemes for transmitting information or data reliably over a **noisy channel** at a fixed transmission rate. The purpose of source coding theory is to show the existence of schemes for compressing data emitted from an **information source** and reproducing them within tolerable limits of distortion. Both theories are based profoundly on the theory of probability, statistics, and the theory of stochastic processes.

B. Entropy

In information theory it is customary to call a finite set $A = \{\alpha_1, \dots, \alpha_n\}$ an **alphabet** and to call its elements the **letters** of the alphabet. The simplest model of information sources consisting of an alphabet A and a \dagger probability distribution p over A is denoted by $X = [A, p]$, which can be regarded as a \dagger finite probability space, where the probability of outcome of a letter $\alpha_i \in A$ from the source is denoted by $p_i = p(\alpha_i)$. A real-valued function defined by $I(\alpha_i) = -\log p(\alpha_i)$ is called the **self-information** of the event $\alpha_i \in A$. The \dagger mean value of the self-information, i.e.,

$$H(X) = \sum_{i=1}^n -p(\alpha_i) \log p(\alpha_i),$$

is called the **entropy** of the source. This can be interpreted as a measure of the average a priori uncertainty as to which letter will emanate from the source, or a measure of the average amount of information one obtains upon receiving a single letter from the source. The unit of information or entropy for base e logarithms is called a **nat**, while that for base 2 logarithms is called a **bit**.

Let $A = \{\alpha_1, \dots, \alpha_n\}$ and $B = \{\beta_1, \dots, \beta_m\}$ be two alphabets. Let $r(\alpha_i, \beta_j)$ be a †joint probability distribution defined on the product AB , and denote the †probability space by $XY = [AB, r]$. Then the joint distribution gives rise to †marginal distributions $p(\alpha_i) = \sum_{j=1}^m r(\alpha_i, \beta_j)$ and $q(\beta_j) = \sum_{i=1}^n r(\alpha_i, \beta_j)$ and to †conditional distributions $p(\alpha_i | \beta_j) = r(\alpha_i, \beta_j)/q(\beta_j)$ and $q(\beta_j | \alpha_i) = r(\alpha_i, \beta_j)/p(\alpha_i)$. Probability spaces $X = [A, p]$ and $Y = [B, q]$ are subspaces of the probability space $XY = [AB, r]$. The entropy of $XY = [AB, r]$ is defined by

$$H(XY) = \sum_{i=1}^n \sum_{j=1}^m -r(\alpha_i, \beta_j) \log r(\alpha_i, \beta_j)$$

as well. A real-valued function $I(\alpha_i | \beta_j) = -\log p(\alpha_i | \beta_j)$ is called the **conditional self-information**. The average of $I(\alpha_i | \beta_j)$ over α_i defined by

$$H(X | \beta_j) = \sum_{i=1}^n p(\alpha_i) I(\alpha_i | \beta_j)$$

is called the **conditional entropy** of X for given $\beta_j \in B$. The average of $H(X | \beta_j)$ over β_j defined by

$$H(X | Y) = \sum_{j=1}^m q(\beta_j) H(X | \beta_j)$$

is called the conditional entropy of X for given Y . The conditional entropy of Y for given X , denoted by $H(Y | X)$, is defined similarly. Then it can be shown that

$$H(XY) = H(Y) + H(X | Y) = H(X) + H(Y | X),$$

$$H(XY) \leq H(X) + H(Y),$$

$$H(X | Y) \leq H(X),$$

$$H(Y | X) \leq H(Y),$$

where equalities in the last three expressions hold if and only if $r(\alpha_i, \beta_j) = p(\alpha_i)q(\beta_j)$ for all $\alpha_i \in A$ and $\beta_j \in B$.

C. Information Sources

Given a finite alphabet A , we consider the finite product set $A^N = A \times \dots \times A$ and the doubly infinite product $A^Z = \prod_{k=-\infty}^{\infty} A_k$, where $A_k = A$, $k = 0, \pm 1, \pm 2, \dots$. Let \mathcal{F}_A be the σ -algebra generated by all †cylinder sets in A^Z . Given a †probability measure P over \mathcal{F}_A , an information source is defined as a probability space $[A^Z, P]$ or as a †random process $\mathbf{X} = \dots X_{-1} X_0 X_1 X_2 \dots$. When P is invariant under the †shift transformation T on A^Z , i.e., $P(E) = P(TE)$ for any $E \in \mathcal{F}_A$, then $[A^Z, P]$ is said to be a **stationary source**. In particular, if $P(E) = 0$ or 1 whenever $TE = E \in \mathcal{F}_A$, then the information source is said to be **ergodic**. If \mathbf{X} is an †independently and identically distributed

random process, the source $\mathbf{X} = [A^Z, P]$ is said to be **memoryless**.

For a given \mathbf{X} , denote the subsequence $X_1 \dots X_N$ of \mathbf{X} by X^N . Then a probability measure P on A^N and a finite probability space $X^N = [A^N, P^N]$ are naturally induced. The entropy of the stationary information source $\mathbf{X} = [A^Z, P]$ is defined as $H(\mathbf{X}) = \lim_{N \rightarrow \infty} H(X^N)/N$ or as $H(\mathbf{X}) = \lim_{N \rightarrow \infty} H(X_N | X^{N-1})$, because both limits exist and are identical. If \mathbf{X} is memoryless, the entropy of \mathbf{X} is equivalent to that of $X_1 = [A^1, P^1]$, i.e., $H(\mathbf{X}) = \sum_{i=1}^n -p_i \log p_i$, where $p_i = p^1(X_1 = \alpha_i)$ and $H(X^N) = NH(X^1) = NH(\mathbf{X})$.

D. Source Coding Theorem

Let $x^N = x_1 \dots x_N$ be a sequence of N consecutive letters from the source $[A^Z, P]$. Suppose that we wish to encode such sequences into fixed-length code words $u^L = u_1 \dots u_L$ consisting of L letters from a code alphabet U of size v . The number $R = (\log v^L)/N$ is called the **coding rate** per source letter. A mapping $\varphi: A^N \rightarrow U^L$ is called an encoder and $\varphi: U^L \rightarrow A^N$ a decoder. The set U^L with specified encoder-decoder pair $[\varphi, \psi]$ is called a **code** with rate $R = (L \log v)/N$. The **error probability** of the code is defined by

$$P_e[\varphi, \psi] = P[\psi\{\varphi(X^N) \neq X^N\}].$$

The fixed-length **source coding theorem** [1, 2] states: Let $\mathbf{X} = [A^Z, P]$ be a stationary ergodic source. Then for any $\delta > 0$, if $R \geq H(\mathbf{X}) + \delta$, there exists a code $[\varphi, \psi]$ with rate R such that the error probability $P_e[\varphi, \psi]$ can be made arbitrarily small by making N sufficiently large. Conversely, if $R \leq H(\mathbf{X}) - \delta$ then for any code with rate R , $P_e[\varphi, \psi]$ must become arbitrarily close to 1 as $N \rightarrow \infty$.

This theorem is trivial if $R \geq \log n$. For memoryless sources, the theorem follows immediately from the †weak law of large numbers, which implies that for arbitrary $\varepsilon > 0$ and $\delta > 0$ there exists an integer $N_0(\varepsilon, \delta)$ such that for all $N > N_0(\varepsilon, \delta)$

$$P\left[\left|\frac{-\log P^N(X^N)}{N} - H(\mathbf{X})\right| > \delta\right] < \varepsilon.$$

The validity of this property for stationary ergodic sources was proved by B. McMillan [2]. In case of memoryless sources, the exact asymptotic form of the error probability for an optimal code with rate R was given by F. Jelinek [3], I. Csiszár and G. Longo [4], Blahut [5], and Longo and A. Sgarro [6] as

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log P_e[\varphi, \psi] = \min_{q: H(q) \geq R} D(q \| p),$$

where p is the source distribution, q denotes a

probability distribution on A , and

$$D(q \| p) = \sum_{i=1}^n q_i \log \frac{q_i}{p_i},$$

which is called **Kullback's discrimination information** [7] or the **divergence**.

Source sequences x^N can be encoded into variable-length code words consisting of letters from alphabet U of size v . A set of n^N code words is called a **prefix condition code** if there is no code word which is equivalent to the prefix of any other code word. Denote the length of the code word corresponding to a source sequence x^N by $L(x^N)$ and the average length of the code words per source letter by

$$\frac{1}{N} \bar{L} = \sum_{x^N} \frac{1}{N} P^N(x^N) L(x^N).$$

The variable length source coding theorem states: Given a memoryless source \mathbf{X} with entropy $H(\mathbf{X})$, there is a prefix condition code such that the average length of the code words per source letter satisfies

$$H(\mathbf{X}) \leq \bar{L} \log v \leq H(\mathbf{X}) + \frac{1}{N} \log v.$$

This theorem is valid for stationary ergodic sources if $(\log v)/N$ in the last term is replaced by $\varepsilon(N)$, where $\varepsilon(N) \rightarrow 0$ as $N \rightarrow \infty$.

These two theorems are referred to as **noiseless source coding theorems**.

E. Source Coding Theorem with Fidelity Criterion

The noiseless source coding theorem implies that the average number of code letters per source letter can be reduced to the source entropy $H(\mathbf{X})$ under the requirement that the source sequence be exactly reproduced from the encoded sequence. If an approximate reproduction of the source sequence to within a given fidelity criterion is required, the coding rate per source letter must be reduced further to a certain value below the source entropy.

Suppose that a **distortion measure** $d(\alpha_i, \alpha_j)$ is defined for $\alpha_i, \alpha_j \in A$, where it is assumed that $d(\alpha_i, \alpha_j) \geq 0$ and $d(\alpha_i, \alpha_i) = 0$. For blocks $x^N = x_1 \dots x_N$ and $y^N = y_1 \dots y_N$, define

$$d_N(x^N, y^N) = \frac{1}{N} \sum_{k=1}^N d(x_k, y_k).$$

Any set $\mathcal{C} = \{y_1^N, \dots, y_M^N\}$, $y_m^N \in A^N$, of reproducing words is called a source code of block length N . Each source sequence x^N of the source $\mathbf{X} = [A^Z, P]$ is mapped into whichever code word $y_m^N \in \mathcal{C}$ minimizes $d_N(x^N, y^N)$, i.e.,

$$d_N(x^N, y_m^N) = \min_{y^N \in \mathcal{C}} d_N(x^N, y^N),$$

and the suffix m is transmitted. Hence the coding rate per source letter is $R = (\log M)/N$, and the average distortion is

$$\bar{d}_N(\mathcal{C}) = \sum_{x^N \in A^N} P^N(x^N) \min_{y^N \in \mathcal{C}} d_N(x^N, y^N).$$

The problem is how far we can reduce the rate under the condition that the average distortion keeps satisfying a given fidelity criterion, which is specified as a maximum tolerable value d for the average distortion.

The **source coding theorem with a fidelity criterion** states: Let $\mathbf{X} = [A^Z, P]$ be a memoryless source. For any specified $d \geq 0$, any $\varepsilon > 0$, and $\delta > 0$, there exists a source code \mathcal{C} with rate $R \geq R(d) + \delta$ and with sufficiently large block length N for which the average distortion satisfies

$$\bar{d}_N(\mathcal{C}) \leq d + \varepsilon,$$

where $R(d)$ is the **rate distortion function** defined by

$$R(d) = \min_{W \in \mathcal{W}(d)} I(p; W),$$

$$I(p; W) = \sum_{i=1}^n \sum_{j=1}^n p_i W(j|i) \log \frac{W(j|i)}{\sum_{i=1}^n p_i W(j|i)},$$

$$\mathcal{W}(d) = \left\{ W: \sum_{i=1}^n \sum_{j=1}^n p_i W(j|i) d(\alpha_i, \alpha_j) \leq d \right\},$$

where $W(j|i)$ denotes a conditional probability distribution referred to as the test channel, and where $I(p; W)$ is called the **mutual information**. It should be noted that $R(0) = H(\mathbf{X})$. The rate distortion function $R(d)$, which was first defined by Shannon [1, 8], is closely related to the ε -**entropy** introduced by A. N. Kolmogorov [9]. $R(d)$ is a monotonically decreasing and convex function.

The theorem was first proved by Shannon [8], and was extended by R. G. Gallager [10] to the case of stationary ergodic sources with discrete alphabets and by T. Berger [11] to stationary ergodic sources with abstract alphabets. More recently, R. G. Gray and L. D. Davisson [12] have proved source coding theorems without the ergodic assumption for stationary sources subject to a fidelity criterion. The proof was based on Rokhlin's ergodic decomposition theorem [13]. Other important source coding theorems have been obtained by F. Jelinek [14] for tree codes, A. J. Viterbi and J. K. Omura [15] for trellis codes, and Gray, D. L. Neuhoff, and D. S. Ornstein [16] for sliding block codes.

The rate distortion function for memoryless Gaussian sources subject to squared error distortion was given by Shannon [8], and that for autoregressive Gaussian sources was determined by Kolmogorov [9], M. S. Pinsker

[17], Berger [11], Gray [18], and T. Hashimoto and S. Arimoto [19].

F. Channel Coding Theory

A mathematical model for a channel over which information is transmitted is specified in terms of the set of possible inputs, the set of outputs, and a probability measure on the output events conditional on each input. The simplest channels are the noiseless ones, for which there is a one-to-one correspondence between input and output and no loss of information in transmission through the channel. The second simplest channels are **discrete memoryless channels** (DMCs) which are defined as follows: The input and the output are sequences of letters from finite alphabets (say, $\alpha_i \in A$ and $\beta_i \in B$), and the output letter at a given time depends statistically only on the corresponding input letter. That is, a DMC is characterized by a fixed conditional probability distribution $W(j|i) = W(\beta_j|\alpha_i)$ because the probability measure on the input and output sequences satisfies

$$W^N(y^N|x^N) = \prod_{k=1}^N W(y_k|x_k), \quad x_k \in A, \quad y_k \in B.$$

By $U = \{1, \dots, M\}$ we denote the set of integers each of which is assigned to each corresponding possible message from the source. A mapping $\varphi: U \rightarrow A^N$ induces a collection $\mathcal{C} = \{x_1^N, \dots, x_M^N\}$, called the **block code** with rate $R = (\log M)/N$; each element is called a code word. Only code words are transmitted over the channel. A mapping $\psi: B^N \rightarrow U$ is called the decoding. Thus, given an encoding and decoding pair $[\varphi, \psi]$, the error probability is defined by

$$P_e[\varphi, \psi] = \frac{1}{N} \sum_{m=1}^M \sum^* W^N(y^N|\varphi(m)),$$

where the summation \sum^* is taken with respect to all y^N such that $\psi(y^N) \neq m$. The **capacity** for a DMC is defined by

$$C = \max_p I(p; W) \\ = \max_p \sum_{i=1}^n \sum_{j=1}^m p_i W(j|i) \log \frac{W(j|i)}{\sum_{i=1}^n p_i W(j|i)}.$$

The fundamental theorem of channel coding theory states: Given a DMC with capacity $C > 0$, there exist a block code with rate R below capacity C , a pair $[\varphi, \psi]$, and a function $E(R) > 0$ for $0 \leq R < C$ such that the error probability satisfies

$$P_e[\varphi, \psi] < \exp\{-NE(R)\}.$$

This was first discovered by Shannon [1], and its precise proof was first given by A.

Feinstein [20]. The precise expression of $E(R)$ was subsequently given by P. Elias [21], R. M. Fano [22], and Gallager [23]. The best expression of $E(R)$, due to Gallager [23], is

$$E(R) = \max\{E_r(R), E_{ex}(R + O(N^{-1}))\},$$

where

$$E_r(R) = \max_{0 \leq \rho \leq 1} \max_p \left[-\rho R - \log \sum_j \left\{ \sum_i p_i W(j|i)^{1/(1+\rho)} \right\}^{1+\rho} \right],$$

$$E_{ex}(R) = \sup_{\rho \geq 1} \max_p \left[-\rho R - \log \sum_{i,k} p_i p_k \left\{ \sum_j \sqrt{W(j|i)W(j|k)} \right\}^{1/\rho} \right]$$

The converse of the fundamental theorem states: Given a DMC with capacity C , for any block code with rate R above C and any pair $[\varphi, \psi]$, the error probability satisfies

$$P_e[\varphi, \psi] > 1 - \exp\{-NE(R)\},$$

where $E(R)$ is a function positive for $R > C$. This was first proved by J. Wolfowitz [24], and the precise expression, found by S. Arimoto [25], is described by

$$E(R) = \max_{-1 \leq \rho \leq 0} \min_p \left[-\rho R - \log \sum_j \left\{ \sum_i p_i W(j|i)^{1/(1+\rho)} \right\}^{1+\rho} \right].$$

The fundamental theorem of coding theory and its converse imply that the capacity is a critical rate; above the capacity, information cannot be transmitted reliably through the channel. Unfortunately, there is, in general, no direct method for computing the capacity, and therefore an iterative method was proposed by Arimoto [26]. Another iterative method for computing the rate distortion function was given by R. E. Blahut [27].

A discrete channel with memory is, in general, defined by a list of probability measure $\{W_x, x \in A^Z\}$ on a \dagger measurable space $\{B^Z, \mathcal{F}_B\}$ such that for each $F \in \mathcal{F}_B$, $W_x(F)$ is a \dagger measurable function of x . A channel W is called stationary if $W_{Tx}(TF) = W_x(F)$ for all $x \in A^Z$ and all $F \in \mathcal{F}_B$. Given an input source $[A^Z, P]$ and a channel W , connecting the input to the channel induces a joint process of input and output, denoted by $[A^Z \times B^Z, PW]$ where PW is a measure on $\{A^Z \times B^Z, \mathcal{F}_{A \times B}\}$. If a joint process $[A^Z \times B^Z, PW]$ is stationary the mutual information between input and output is defined by

$$I(X; Y) = H(X) + H(Y) - H(XY).$$

A channel W is called ergodic if for every

ergodic input $[A^Z, P]$ the induced joint process $[A^Z \times B^Z, PW]$ is ergodic. The channel capacity for a discrete stationary channel is defined in various ways. The important ones are the **ergodic capacity** defined by $C_e = \sup.$ of $I(X; Y)$ with respect to all stationary ergodic sources $X = [A^Z, P]$, and the **stationary capacity** defined by $C_s = \sup.$ of $I(X; Y)$ w.r.t. all stationary sources. K. R. Parthasarathy [28] proved $C_e = C_s$ for all discrete stationary channels. In another way J. Nedoma [29] introduced the operational source/channel block coding capacity C_{scb} , which is defined as the supremum of the entropies of all admissible stationary ergodic sources in the sense that there exist source/channel block codes such that the error probability $P_e \rightarrow 0$ as block length $N \rightarrow \infty$. Nedoma [29] also pointed out an example of a stationary channel where $C_s > C_{scb}$. Hence **block coding theorems** have been proved for various channels: for **finite memory channels** by A. I. Khinchin [30], K. Takano [31], Nedoma [29], and Feinstein [32], for **\bar{d} -continuous channels** by Gray and Ornstein [33], and for **almost finite memory channels** by D. L. Neuhoff and P. C. Shields [34].

References

- [1] C. E. Shannon, A mathematical theory of communication, Bell System Tech. J., 27 (1948), 379–423, 623–656.
- [2] B. McMillan, The basic theorems of information theory, Ann. Math. Statist., 24 (1953), 196–219.
- [3] F. Jelinek, Probabilistic information theory, McGraw-Hill, 1968.
- [4] I. Csiszár and G. Longo, On the error exponent for source coding, Studia Math. Acad. Hung., 6 (1971), 181–191.
- [5] R. E. Blahut, Hypothesis testing and information theory, IEEE Trans. Information Theory, 20 (1974), 405–417.
- [6] G. Longo and A. Sgarro, The source coding theorem revisited: A combinatorial approach, IEEE Trans. Information Theory, 25 (1979), 544–548.
- [7] S. Kullback, Information theory and statistics, Wiley, 1957.
- [8] C. E. Shannon, Coding theorems for a discrete source with a fidelity criterion, IRE National Convention Record, pt. 4 (1959), 142–163.
- [9] A. N. Kolmogorov, On the Shannon theory of information transmission in the case of continuous signals, IRE Trans. Information Theory, 2 (1956), 102–108.
- [10] R. G. Gallager, Information theory and reliable communication theory, Wiley, 1968.
- [11] T. Berger, Rate distortion theory for sources with abstract alphabets and memory, Information and Control, 13 (1968), 254–273.
- [12] R. G. Gray and L. D. Davisson, Source coding theorems without the ergodic assumption, IEEE Trans. Information Theory, 20 (1974), 502–516.
- [13] V. A. Rokhlin, On the fundamental ideas of measure theory, Amer. Math. Soc. Transl., 71 (1952).
- [14] F. Jelinek, Tree-encoding of memoryless time-discrete sources with a fidelity criterion, IEEE Trans. Information Theory, 15 (1969), 584–590.
- [15] A. J. Viterbi and J. K. Omura, Trellis encoding of memoryless discrete-time sources with a fidelity criterion, IEEE Trans. Information Theory, 20 (1974), 325–332.
- [16] R. M. Gray, D. L. Neuhoff, and D. S. Ornstein, Nonblock source coding with a fidelity criterion, Ann. Probability, 3 (1975), 478–491.
- [17] M. S. Pinsker, Information and information stability of random variables and processes, Holden-Day, 1965.
- [18] R. M. Gray, Information rates of autoregressive processes, IEEE Trans. Information Theory, 16 (1970), 412–421.
- [19] T. Hashimoto and S. Arimoto, On the rate-distortion function for the nonstationary Gaussian AR process, IEEE Trans. Information Theory, 26 (1980), 478–480.
- [20] A. Feinstein, A new basic theorem of information theory, IRE Trans. Information Theory, 4 (1954), 2–22.
- [21] P. Elias, Coding for noisy channels, IRE National Convention Record, pt. 4 (1955), 37–46.
- [22] R. M. Fano, Transmission of information, MIT Press and Wiley, 1961.
- [23] R. G. Gallager, A simple derivation of the coding theorem and some applications, IEEE Trans. Information Theory, 11 (1965), 3–18.
- [24] J. Wolfowitz, The coding of messages subject to chance errors, Illinois J. Math., 1 (1957), 591–606.
- [25] S. Arimoto, On the converse to the coding theorem for discrete memoryless channels, IEEE Trans. Information Theory, 19 (1973), 357–359.
- [26] S. Arimoto, An algorithm for computing the capacity of arbitrary discrete memoryless channels, IEEE Trans. Information Theory, 18 (1972), 14–20.
- [27] R. E. Blahut, Computation of channel capacity and rate-distortion functions, IEEE Trans. Information Theory, 18 (1972), 460–473.
- [28] K. R. Parthasarathy, On the integral representation of the rate of transmission of a stationary channel, Illinois J. Math., 2 (1961), 299–305.

- [29] J. Nedoma, The capacity of a discrete channel, Trans. 1st Prague Conf. Information Theory (1957), 143–181.
- [30] A. Ya. Khinchin, Mathematical foundations of information theory, Dover, 1957.
- [31] K. Takano, On the basic theorems of information theory, Ann. Inst. Statist. Math., 9 (1957), 53–77.
- [32] A. Feinstein, On the coding theorem and its converse for finite-memory channels, Information and Control, 2 (1959), 25–44.
- [33] R. M. Gray and D. S. Ornstein, Block coding for discrete stationary \bar{d} -continuous noisy channels, IEEE Trans. Information Theory, 25 (1979), 292–306.
- [34] D. L. Neuhoff and P. C. Shields, Channels with almost finite memory, IEEE Trans. Information Theory, 25 (1979), 440–447.

214 (XVIII.18) Insurance Mathematics

A. General Remarks

Insurance is a system in which a large number of people contribute a small precalculated amount of money (called a **premium**) to fill the economic need that arises when a person meets adversity. The amount of economic need filled by this system is called the **amount of insurance** (or **amount insured**). The **insurer** is the one who implements the system. **Actuarial mathematics** is the branch of applied mathematics that studies the mathematical basis of insurance, one of the first cases in which mathematics was successfully applied to a social question. Actuarial mathematics can be divided into two branches according to its application. The first includes the calculation of various values of each individual policy, such as premiums or reserves. The second is mainly connected with management of an insurance business and includes the study of reinsurance systems, of the maximum amount of insurance, of the contingency fund, or the analysis of profits. There is only one basic principle in actuarial mathematics, called the **principle of equivalence**. It determines the premium and reserve in each year so that the present value of future premium income of the insurer is equal to the present value of future benefits for each policy.

The basic factors of actuarial calculations are (1) probabilities of contingencies, (2) an expected rate of interest in the future (often referred to as the **assumed rate of interest**), and (3) cost of administration of the system.

Premiums are calculated using these factors and the principle of equivalence. The following is an example of the classical method of calculation for a life insurance policy with the use of “commutation symbols,” which is an old device for the convenience of calculations.

We write P for the net premium (in which the cost of administration is disregarded), P' for the gross premium, T_t for the amount of death benefits payable in the t th year after the policy is issued, E_t for the amount of survival benefits payable at the beginning of the t th year, n for the period for which the insurance is effective, and m for the period for which premiums are to be paid. Let α , β , and γ stand for three positive constants determining the initial expenses $= \alpha(T_1 \text{ or } E_1)$, the premium collection expenses $= \beta P'$, and the general expenses for maintenance $= \gamma(T_t \text{ or } E_t)$. The factor that comes into consideration next is a **model of human death and survival** (measurement).

Assume that l_x is the number of lives attaining age x , and write q_x for the probability that a life of x years will end within one year. Then d_x , the number of lives ending within one year out of l_x , is $l_x q_x$, and l_{x+1} , the number of lives remaining after one year at age $x+1$, is $l_x - d_x = l_x(1 - q_x)$. The commutation symbols commonly employed are defined as follows: Write $v = 1/(1+i)$, where i is the assumed rate of interest; then

$$D_x = l_x v^x, \quad C_x = d_x v^{x+1},$$

$$N_x = \sum_{t=0}^n D_{x+t}, \quad M_x = \sum_{t=0}^n C_{x+t}.$$

For a policy issued at an insured person's age x , the present value of the insurer's future income can be expressed as $P'(N_x - N_{x+m})/D_x$, and the present value of his future payments can be expressed as

$$\frac{1}{D_x} \left(\sum_{t=1}^n T_t C_{x+t-1} + \sum_{t=1}^{n+1} E_t D_{x+t-1} + \alpha(T_1 \text{ or } E_1) D_x + \gamma \sum_{t=0}^n (T_t \text{ or } E_t) D_{x+t-1} + \beta P'(N_x - N_{x+m}) \right).$$

By assuming that the present value of the future income is equal to the present value of the future payments, the value of the gross premium P' is obtained. (The P' obtained from the assumption $\alpha = \beta = \gamma = 0$ is denoted by P and is called the **net premium**. The difference $P' - P$ is called the **loading**.) For a policy in which benefits are payable on disability or contingencies other than death, we have only to obtain a model of contingencies and apply a similar calculation.

B. Liability Reserve

During the term of an insurance contract, it often happens that the present value of the future income is less than the present value of the future payments. If this is the case, the difference is to be held by the insurer as the **liability reserve**. The source of this fund is the past premium income plus interest. The net premium reserve, which disregards expenses, is calculated as

$$V_t = \frac{1}{D_{x+t}} \left(\sum_{r=t+1}^n T_r C_{x+r-1} + \sum_{r=t+1}^{n+1} E_r D_{x+r-1} - P(N_{x+1} - N_{x+m}) \right).$$

Between the net premium P and the net premium reserve V , we have the relation

$$P = (vV_t - V_{t-1}) + (T_t - V_t)vq_{x+t} + E_t.$$

The first term of the right-hand side of this formula is called the **savings premium**, since it is the amount left out of the premium income of the t th year and added to the reserve. The second term is called the **cost of insurance** or **risk premium** and is applied to cover the difference between the amount of insurance and that of the existing reserve in case the contingency of death arises. The third term is applied to the payment of the survival benefits (or annuities in case of an **annuity contract**). If $T_t - V_t$, the amount of risk insured by the insurer, is positive for all values of t during the period of insurance, the policy is called **death insurance**. On the other hand, if the value of $T_t - V_t$ is negative for all values of t , the policy is called **survival insurance**. If the value of $T_t - V_t$ varies between positive and negative according to the different values of t , the policy is called **mixed insurance**. If T_t is always equal to V_t , the policy constitutes mere savings. Most of the insurance policies issued today are one or another type of death insurance, while life annuity policies are a type of survival insurance. For a long time, studies have been made of the effect on premiums and reserves of changes in the three basic factors (1), (2), and (3) in Section A.

C. Risk Theory

Risk theory occupies a special position in the field of actuarial mathematics. Actuarial mathematics was first born from the theory of probability. Since the modern theory of probability based on measure theory was developed by A. N. Kolmogorov and other mathematicians (\rightarrow 342 Probability Theory), new ap-

proaches have inevitably been made to actuarial mathematics. An outstanding example is risk theory.

Risk theory can be divided into two branches. One is called **classical risk theory** (or **individual risk theory**), in which the profit or loss that may result during a certain term of an insurance contract is regarded as a \dagger random variable. Since the insurer's profit equals the sum of these random variables over all the individual contracts, various probability functions can be obtained by applying the theory of probability. The second, called **collective risk theory**, pays no attention to each individual contract but studies changes in the insurer's balance as a whole with the lapse of time. The basis of collective risk theory was given by F. Lundberg, H. Cramér, and other mathematicians.

We explain the collective risk theory following Cramér [5]. For simplicity we consider an insurer who issues no policies other than death insurance and makes no expenditures except the policy claims. Suppose that during the time interval $(t, t + \Delta t)$ contingency occurs with probability $\lambda \Delta t + o(\Delta t)$ independently of the past. We denote by $F(x)$ the \dagger distribution function of the amount to be paid by the insurer when a contingency occurs. Then the number of contingencies in $(0, t)$, $N(t)$, is a \dagger Poisson process with parameter λ , and the total expenditure of the insurer during the time interval $(0, t)$, $X(t)$, is a \dagger compound Poisson process (\rightarrow 5 Additive Processes) such that

$$E(e^{izX(t)}) = \exp \left(\lambda t \int_0^\infty (e^{izx} - 1) dF(x) \right).$$

If p is the premium income per unit time, then we have $p = \lambda \int_0^\infty x dF(x)$, because $E(X(t)) = pt$ holds by the principle of equivalence. If u denotes the initial fund of the insurer, then the fund reserved at time t , $Y(t)$, equals $u + pt - X(t)$. $Y(t)$ is a \dagger Lévy process (\rightarrow 5 Additive Processes) such that

$$\begin{aligned} E(e^{izY(t)}) \\ = \exp \left(izu + \lambda t \int_0^\infty (e^{izx} - 1 - izx) dF(x) \right). \end{aligned}$$

The probability that $Y(t) < 0$ for some time $t < \infty$ is called the **ruin probability**, which depends on the initial fund u , denoted by $\psi(u)$. This satisfies the following \dagger integral equation of Volterra type:

$$\begin{aligned} \frac{p}{\lambda} \psi(u) = \int_u^\infty Q(x) dx + \int_0^u \psi(x) Q(u-x) dx \\ (Q(x) = 1 - F(x)), \end{aligned}$$

from which one can derive an asymptotic relation $\psi(u) \sim Ce^{-Ru}$ as $u \rightarrow \infty$, where R and C

are positive constants depending only on λ and F . See [6] for recent developments in risk theory.

References

- [1] C. W. Jordan, Society of Actuaries' textbook on life contingencies, Society of Actuaries, 1952.
- [2] W. Saxer, Versicherungsmathematik, Springer, I, 1955; II, 1958.
- [3] E. Zwinggi, Versicherungsmathematik, Birkhäuser, 1945.
- [4] P. F. Hooker and L. H. Longley-Cook, Life and other contingencies I, Cambridge Univ. Press, 1953.
- [5] H. Cramér, Collective risk theory, Nordiska Bokhandeln, 1955.
- [6] H. Bühlmann, Mathematical methods in risk theory, Springer, 1970.

215 (XIX.6) Integer Programming

A. General Remarks

Integer programming, in its broadest sense, addresses itself to either minimization or maximization of a functional f over some discrete set S in \mathbf{R}^n , but it is usually understood as dealing with questions related to linear programming problems (\rightarrow 255 Linear Programming) with additional integrality conditions on the variables, namely, the problem P_0 : Minimize $\{c'x \mid Ax = b, x \geq 0, x_j \text{ an integer}, j = 1, \dots, n_1 (\leq n)\}$, where $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $c \in \mathbf{R}^n$ are given data and $x = (x_1, \dots, x_n)' \in \mathbf{R}^n$ is a vector. P_0 is called a **pure integer programming problem** or an **all-integer programming problem** if $n = n_1$, and a **mixed integer programming problem** if $n_1 < n$. In particular, P_0 is called a **0-1 integer programming problem** if all the integer variables are restricted to be equal to either 0 or 1. We write

$$X^0 = \{x \in \mathbf{R}^n \mid Ax = b, x_j \geq 0, j = 1, \dots, n\},$$

$$X^I = X^0 \cap \{x \in \mathbf{R}^n \mid x_j \text{ is an integer}, j = 1, \dots, n_1\},$$

and assume for simplicity that (i) $X^I \neq \emptyset$, (ii) X^0 is bounded, and (iii) all the components of A and b are integers. P_0 arises not only as a mathematical model for an optimization problem where some of the decision variables have indivisible minimum units but also as one for many optimization problems with some logical and/or combinatorial constraints [1].

Algorithms for solving P_0 can be classified into algebraic methods and enumerative methods. Both are based upon linear programming and **relaxation** techniques [2] in which some of the constraints of P_0 are temporarily relaxed (\rightarrow 264 Mathematical Programming).

B. Cutting Plane Methods

The origin of this class of algorithms is the **fractional cutting plane algorithm** proposed in 1958 by R. E. Gomory [3], the outline of which is described as:

- (1) Let $X = X^0$.
- (2) Solve a linear programming problem: Minimize $\{c'x \mid x \in X\}$, and let \bar{x} be its \dagger optimal solution. If the \bar{x}_j are integer for all $j \leq n_1$, then stop (\bar{x} is optimal to P_0). Otherwise, go to (3).
- (3) Generate a half-space $H = \{x \in \mathbf{R}^n \mid \pi'x \geq \pi_0\}$ ($\pi \in \mathbf{R}^n, \pi_0 \in \mathbf{R}^1$), where H (i) contains X^I and (ii) does not contain \bar{x} . Return to (2) by replacing X by $X \cap H$.

Here, a linear programming problem obtained by relaxing integrality conditions is solved, and as long as $\bar{x} \notin X^I$, an inequality satisfying the two conditions (i), (ii) of (3) is introduced. Such an inequality $\pi'x \geq \pi_0$ or equality $\pi'x = \pi_0$ is called a **cut** or a **cutting plane**. Gomory devised a **Gomory cut** using (relaxed) integrality conditions on the variables, and showed that the algorithm above produces a point of X^I in finitely many steps. Some of the other algorithms using cutting planes are the **all-integer algorithm** (Gomory, 1963) and the **primal all-integer algorithm** (R. D. Young, *Operations Res.*, 16 (1968)). These algorithms, however, are generally slow and behave erratically, so that it is believed that they cannot in practice serve as general-purpose algorithms.

C. Other Algebraic Methods

Gomory [4], again in 1965, proposed a **group-theoretic approach** to P_0 . This method is based upon the following observation: Let x_B be the vector of \dagger basic variables (\rightarrow 255 Linear Programming) associated with a \dagger dual feasible basis of a linear programming problem \bar{P}_0 : Minimize $\{c'x \mid x \in X^0\}$, and let \tilde{X}^I be the set generated from X^I by relaxing the nonnegativity constraints on x_B . Then \tilde{X}^I can be shown to have a \dagger cyclic group structure, so that a **group minimization problem** P_G : minimize $\{c'x \mid x \in \tilde{X}^I\}$ can be solved as a \dagger shortest path problem on a directed graph with a special structure (\rightarrow 186 Graph Theory). If the optimal solution \tilde{x} of P_G satisfies $\tilde{x}_B \geq 0$, then it is optimal for P_0 . If, on the other hand, $\tilde{x}_B \not\geq 0$,

then a branch and bound algorithm described later can be applied for a systematic search of integer points near \tilde{x} . This algorithm is reported to produce good results when the size of the associated graph is not excessively large (G. A. Gorry et al., *Management Sci.*, 17 (1971)). Some of the other important results in this area are: (i) the theory of **subadditive cuts** (R. Gomory et al., *Math. Prog.*, 3 (1972)) and **disjunctive cuts** (E. Balas, in *Nonlinear Programming 2*, O. Mangasarian et al. (eds.), Academic Press, 1975), (ii) research on facial structures of the **integer polyhedron** $\text{co}X^I$ (convex hull of X^I) for some of the more important integer programming problems, such as that of the **corner polyhedron** $\text{co}\tilde{X}^I$ [5], knapsack polytopes (E. Balas, *Math. Prog.*, 8 (1975)), and traveling-salesman polytopes (M. Grötschel et al., *Math. Prog.*, 16 (1979)). It should be pointed out, however, that more difficulties of P_0 have been revealed rather than resolved through the intensive research in this area. Incidentally, the 0-1 integer programming problem is known to be ${}^{\dagger}NP$ -complete (\rightarrow 71 Complexity of Computations).

D. Enumerative Methods

Another large class of algorithms for solving P_0 consists of the **branch and bound methods**, first proposed by Land and Doig [6] in 1960. An outline of the improved version (R. J. Dakin, *Computer J.*, 8 (1965)) is as follows.

- (1) Let $\mathcal{P} = \{P_0\}$, $z^* = \infty$, $x^* = \text{undefined}$.
- (2) If $\mathcal{P} = \emptyset$, then stop (x^* is optimal to P_0). Otherwise, choose from \mathcal{P} the problem P_i : Minimize $\{c'x \mid x \in X_i^I\}$.
- (3) Solve \bar{P}_i : Minimize $\{c'x \mid x \in X_i\}$, in which the integrality condition is relaxed from the constraints of P_i . If \bar{P}_i has an optimal solution \tilde{x} , then go to (4). Otherwise, return to (2).
- (4) If $\tilde{x} \in X_i^I$ and $c'\tilde{x} < z^*$, then let $\tilde{x} \rightarrow x^*$, $c'\tilde{x} \rightarrow z^*$, and return to (2). If $\tilde{x} \in X_i^I$ and $c'\tilde{x} \geq z^*$, then return to (2). Otherwise, go to (5).
- (5) Choose $j \leq n_1$, for which \tilde{x}_j is not integral and generate the two **subproblems** P_j^+ : minimize $\{c'x \mid x \in X_i^I, x_j \geq [\tilde{x}_j] + 1\}$, and P_j^- : minimize $\{c'x \mid x \in X_i^I, x_j \leq [\tilde{x}_j]\}$, in both of which $[\tilde{x}_j]$ represents the largest integer not exceeding \tilde{x}_j . Let $\mathcal{P} \cup \{P_j^-, P_j^+\} \rightarrow \mathcal{P}$ and return to (2).

The best point of X^I identified during the preceding steps is denoted by x^* and called an incumbent. In summary, the branch and bound method chooses one subproblem P_i from the problem list \mathcal{P} and estimates the lower bound of its optimal objective functional value. If the lower bound is worse than the current incumbent, then P_i is discarded, whereas P_i is separated into two subproblems if no conclusion can be reached. This process

is continued until the problem list \mathcal{P} becomes empty, thereby implicitly checking all points of X^I . The above method is called an LP-based branch and bound method because linear programming techniques are employed to obtain a lower bound. The branch and bound method tends to require a large amount of storage, but many engineering improvements on the method of choosing (i) a subproblem P_i and (ii) a branching variable x_j , and several improvements of the bounding techniques, in addition to the substantial progress made in linear programming codes, enable us to solve a problem of size $n \approx 100$. In particular, an improved version of the **implicit enumeration method**, proposed by E. Balas [7] for 0-1 integer programming problems which uses logical conditions for obtaining lower bounds, is known to be able to solve rather large 0-1 integer programming problems (A. Geoffrion, *Operations Res.*, 17 (1969)).

E. Other Topics

The **partitioning algorithm** [8], in which integer variables are varied parametrically, is reported to work well for 0-1 mixed integer problems with relatively few integer variables. As people begin to realize the intrinsic difficulty of P_0 , they pay more attention to **heuristic algorithms** or approximate algorithms to obtain a good but not necessarily optimal solution. Among heuristic methods, the interior path method [9], which elaborates simple ideas such as rounding of the optimal solution of \bar{P}_0 , has been reported to work well for problems in which X^0 has a nonempty interior. Also, more emphasis is being placed on special purpose algorithms for solving practical problems, such as ${}^{\dagger}\text{set partitioning}$ and the ${}^{\dagger}\text{traveling salesman problem}$, etc. [1]. P_0 is a typical nonconvex programming problem, and no practically useful ${}^{\dagger}\text{duality theorem}$ is available. Hence it is difficult to perform sensitivity and/or post-optimality analysis. Some research in this area has emerged recently (e.g., C. J. Piper et al., *Management Sci.*, 22 (1976)), but it looks as if it will be several years before a reasonably good procedure becomes available.

References

- [1] R. S. Garfinkel and G. L. Nemhauser, *Integer programming*, Wiley, 1972.
- [2] A. M. Geoffrion and R. E. Marten, *Integer programming: A framework and state-of-the-art survey*, *Management Sci.*, 18 (1972), 465–491.

- [3] R. E. Gomory, Outline of an algorithm for integer solutions to linear programs, *Bull. Amer. Math. Soc.*, 64 (1958), 275–278.
- [4] R. E. Gomory, On the relation between integer and non-integer solutions to linear programs, *Proc. Nat. Acad. Sci. US*, 53 (1965), 260–265.
- [5] R. E. Gomory, Some polyhedra related to combinatorial problems, *Linear Algebra and Appl.*, 2 (1969), 451–558.
- [6] A. H. Land and A. G. Doig, An automatic method for solving discrete programming problems, *Econometrica*, 28 (1960), 497–520.
- [7] E. Balas, An additive algorithm for solving linear programs with zero-one variables, *Operations Res.*, 13 (1965), 517–546.
- [8] J. F. Benders, Partitioning procedures for solving mixed-variables programming problems, *Numer. Math.*, 4 (1962), 238–252.
- [9] F. S. Hillier, Efficient heuristic procedures for integer linear programming with an interior, *Operations Res.*, 17 (1969), 600–637.

216 (X.10) Integral Calculus

A. The Riemann Integral

Let $f(x)$ be a bounded real-valued function defined on an interval $[a, b]$. We shall divide this interval $I = [a, b]$ into subintervals $I_i = [x_{i-1}, x_i]$ ($i = 1, \dots, n$) by a finite number of points x_i ($a = x_0 < x_1 < \dots < x_n = b$). This division into subintervals is uniquely determined by the set $D = \{x_i\}$, called the **partition** of I . We set $M_i = \sup_{x \in I_i} f(x)$, $m_i = \inf_{x \in I_i} f(x)$, and put $\bar{\sigma}(D) = \sum_{i=1}^n M_i(x_i - x_{i-1})$, $\underline{\sigma}(D) = \sum_{i=1}^n m_i(x_i - x_{i-1})$. Considering all possible partitions D of I , we set $\int_a^b f(x) dx = \inf_D \bar{\sigma}(D)$, $\int_a^b f(x) dx = \sup_D \underline{\sigma}(D)$, which are called the **Riemann upper integral** and **Riemann lower integral** of f , respectively. If they coincide, then the common value is called the **Riemann integral** of f on $[a, b]$ and is denoted by $\int_a^b f(x) dx$. In this case, we say that f is **Riemann integrable** (or simply **integrable**) on $[a, b]$ and call f the **integrand**; a and b are called the **lower limit** and the **upper limit**, respectively. In this case, by **integrating** f from a to b we mean the process of obtaining the value $\int_a^b f(x) dx$.

Darboux's theorem: For each $\varepsilon > 0$ there exists a positive δ such that the inequalities

$$\left| \bar{\sigma}(D) - \int_a^b f(x) dx \right| < \varepsilon,$$

$$\left| \underline{\sigma}(D) - \int_a^b f(x) dx \right| < \varepsilon$$

hold for any partition D with $\delta(D) = \max_i (x_i - x_{i-1}) < \delta$. In other words, we have

$$\lim_{\delta(D) \rightarrow 0} \bar{\sigma}(D) = \int_a^b f(x) dx,$$

$$\lim_{\delta(D) \rightarrow 0} \underline{\sigma}(D) = \int_a^b f(x) dx.$$

From Darboux's theorem it follows that a necessary and sufficient condition for $f(x)$ to be integrable on $[a, b]$ is that for each positive ε there exist a positive δ such that $\delta(D) < \delta$ implies $\bar{\sigma}(D) - \underline{\sigma}(D) = \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) < \varepsilon$.

We call $M_i - m_i$ the **oscillation** of f on I_i and $\bar{\sigma}(D)$ and $\underline{\sigma}(D)$ the **Darboux sums**. Obviously, if f is integrable on $[a, b]$, then for each positive ε there exists a positive δ such that the following inequality holds for every partition $D = \{x_j\}$ with $\delta(D) < \delta$ and for every set of points $\xi_j \in I_j$ ($j = 1, \dots, n$):

$$\left| \sum_{j=1}^n f(\xi_j)(x_j - x_{j-1}) - \int_a^b f(x) dx \right| < \varepsilon.$$

The sum $\sum_{j=1}^n f(\xi_j)(x_j - x_{j-1})$ is often called a **Riemann sum** (or **sum of products**). A function that is continuous on $[a, b]$, or bounded and continuous except for a finite number of points in the interval, is integrable. Furthermore, a bounded function that is continuous on $[a, b]$ except for an infinite number of points x_λ is integrable if for an arbitrary positive number ε there exist a finite number of intervals I_i of which the total length is less than ε and if the set $\{x_\lambda\}$ of exceptional points is contained in $\bigcup I_i$. Generally, a necessary and sufficient condition for a bounded function defined on $[a, b]$ to be integrable is that the set of points where the function is not continuous be of \dagger measure 0 (in the sense of Lebesgue). A function that is either \dagger monotonic on $[a, b]$ (and consequently bounded) or of \dagger bounded variation is integrable. A function that is integrable on $[a, b]$ is integrable on any subinterval of $[a, b]$, the integrand being the restriction of the given function to this subinterval.

B. Basic Properties of Integrals

Let \mathbf{I} be the set of all functions integrable on $[a, b]$. If $f, g \in \mathbf{I}$, then for any numbers α, β , we have $\alpha f + \beta g \in \mathbf{I}$, $fg \in \mathbf{I}$, $\min\{f, g\} \in \mathbf{I}$, $\max\{f, g\} \in \mathbf{I}$, and $f/g \in \mathbf{I}$ provided that there exists a positive constant A such that the inequality $|g| > A$ holds. Furthermore, if $f \in \mathbf{I}$, then $|f| \in \mathbf{I}$; and if $f_n \in \mathbf{I}$ ($n = 1, 2, \dots$) and f_n converges \dagger uniformly to f , then $f \in \mathbf{I}$. Corresponding to these properties, the following formulas hold.

(1) Linearity:

$$\begin{aligned} & \int_a^b (\alpha f(x) + \beta g(x)) dx \\ &= \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx, \end{aligned}$$

where α, β are constants.

(2) Monotonicity: If $f(x) \geq 0$, then

$$\int_a^b f(x) dx \geq 0.$$

If, further, f is continuous at a point $x_0 \in [a, b]$ and $f(x_0) > 0$, then $\int_a^b f(x) dx > 0$.

(3) Additivity with respect to intervals: If a, b , and c are points belonging to an interval on which f is integrable and $a < c < b$, then

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Adopting the conventions that $\int_a^a f(x) dx = 0$ and $\int_b^a f(x) dx = -\int_a^b f(x) dx$, the additivity formula holds independently of the order of a, b , and c .

It follows from (2) that $|\int_a^b f(x) dx| \leq \int_a^b |f(x)| dx$ if $a < b$. Further, if $f_n(x)$ converges to $f(x)$ uniformly on $[a, b]$, then

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx.$$

Replacing $f_n(x)$ by partial sums of a series, we obtain the following theorem: Let $\sum a_n(x)$ be a series in which each term $a_n(x)$ is integrable on an interval $[a, b]$. If the series converges uniformly on $[a, b]$, then the sum $s(x)$ is integrable on $[a, b]$, and the series is **termwise integrable**, that is,

$$\int_a^b s(x) dx = \sum_{n=1}^{\infty} \int_a^b a_n(x) dx.$$

Also, the series $\sum_{n=1}^{\infty} \int_a^x a_n(t) dt$ converges uniformly on $[a, b]$ to the integral $\int_a^x s(t) dt$. Assume that $\sum a_n(x)$ is convergent but not uniformly convergent. If all $a_n(x)$, together with $s(x) = \sum a_n(x)$, are integrable and there is a constant M independent of n such that $|s_n(x)| \leq M$ ($x \in [a, b]$) for all n , where $s_n(x)$ are partial sums, then the series is termwise integrable (C. Arzelà).

The first mean value theorem: If $f(x)$ is continuous on $[a, b]$ and $\varphi(x)$ is integrable and of constant sign on $[a, b]$, then there exists θ ($0 < \theta < 1$) such that

$$\int_a^b f(x) \varphi(x) dx = f(a + \theta(b-a)) \int_a^b \varphi(x) dx.$$

When $\varphi(x) = 1$, we have

$$\int_a^b f(x) dx = f(a + \theta(b-a))(b-a).$$

The second mean value theorem: If $f(x)$ is a positive, monotone decreasing function defined on $[a, b]$ and $\varphi(x)$ is an integrable function, then there exists η ($a < \eta \leq b$) such that

$$\int_a^b f(x) \varphi(x) dx = f(a + 0) \int_a^{\eta} \varphi(x) dx.$$

In the hypothesis of the second mean value theorem, if $f(x)$ is assumed to be monotonic but not necessarily positive, then there exists η ($a \leq \eta \leq b$) such that

$$\begin{aligned} & \int_a^b f(x) \varphi(x) dx \\ &= f(a + 0) \int_a^{\eta} \varphi(x) dx + f(b - 0) \int_{\eta}^b \varphi(x) dx. \end{aligned}$$

H. Okamura (1947) proved that the condition $a \leq \eta \leq b$ can be replaced by $a < \eta < b$.

In the case $f(x) \geq 0$ on $[a, b]$, we consider the figure F bounded by the graph of $f(x)$, the x -axis, and the lines $x = a$ and $x = b$. Then $\bar{\sigma}(D)$ and $\sigma(D)$ are areas of polygons of which one encloses F and the other is enclosed by F , as shown in Fig. 1. Hence it can be shown that the integrability of $f(x)$ in the sense of Riemann is equivalent to the measurability of F in the sense of Jordan. The Riemann integral $\int_a^b f(x) dx$ is the area of F with respect to its [†]Jordan measure.

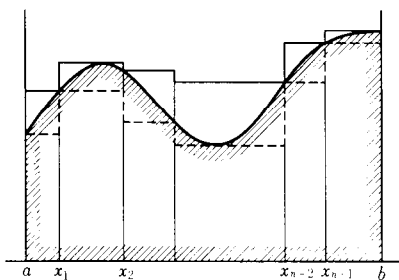


Fig. 1

C. Relation between Differentiation and Integration

Suppose that $f(x)$ is integrable on an interval I . We fix a point a of I and consider the integration $F(x) = \int_a^x f(t) dt$, where x varies in I . The function $F(x)$ is called the **indefinite integral** of $f(x)$. In contrast with this, the integral on a fixed interval, as considered in the previous sections, is often called the **definite integral**. The indefinite integral $F(x)$ is continuous on the interval I and of bounded variation. If $f(x)$ is continuous at a point x_0 in I , then $F(x)$ is differentiable at x_0 and $F'(x_0) = f(x_0)$. In general, if a function $G(x)$ satisfies $G'(x) = f(x)$ everywhere in I , then $G(x)$ is called a **primitive function** of $f(x)$. If $f(x)$ is continu-

ous, the indefinite integral of $f(x)$ is one of the primitive functions of $f(x)$. Furthermore, if a function $G(x)$ is a primitive function of $f(x)$, then any other primitive function can be written in the form $G(x) + C$, where C is a constant, called an **integral constant**. For a continuous function $f(x)$ on $[a, b]$ and any one of its primitive functions $G(x)$, we have

$$\int_a^b f(x) dx = G(b) - G(a) = [G(x)]_a^b$$

(fundamental theorem of calculus) (\rightarrow Appendix A, Table 9). From the differentiation formulas we obtain the following integration formulas:

Integration by parts: If $f(x)$ and $g(x)$ have continuous derivatives on $[a, b]$, then

$$\int_a^b f(x)g'(x) dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx.$$

More generally, if $f(x)$ and $g(x)$ are integrable on $[a, b]$, then

$$\begin{aligned} \int_a^b g(x) \left(\int_a^x f(t) dt \right) dx \\ = \int_a^b f(x) dx \int_a^b g(x) dx - \int_a^b f(x) \left(\int_a^x g(t) dt \right) dx. \end{aligned}$$

Change of variables: If $f(x)$ is integrable on $[a, b]$ and $x = \varphi(t)$ and $\varphi'(t)$ are continuous on $[\alpha, \beta]$, where $a = \varphi(\alpha)$, $b = \varphi(\beta)$ ($a \leq \varphi(t) \leq b$), then

$$\int_a^b f(x) dx = \int_\alpha^\beta f(\varphi(t)) \varphi'(t) dt.$$

D. Improper Integrals

The concept of the integral can be generalized to the case where the integrand or the interval on which integration is accomplished is not bounded. Assume that $f(x)$ is not bounded on $[a, b]$ but is bounded and integrable on any interval $[a, b - \varepsilon]$ ($\subset [a, b]$). If $\int_{a-\varepsilon}^b f(x) dx$ has a finite limit for $\varepsilon \rightarrow 0$, the limit is denoted by $\int_a^b f(x) dx$ and is called the **improper Riemann integral** (or simply **improper integral**) of $f(x)$ on $[a, b]$. For example, if $f(x)$ is continuous on $[a, b]$ and $f(x) = O((b-x)^\alpha)$ for some α ($0 > \alpha > -1$), where O is the \dagger Landau symbol, then the improper integral $\int_a^b f(x) dx$ exists. On the other hand, if f is integrable on $[a + \varepsilon, b]$ for each $\varepsilon > 0$ but not bounded in any neighborhood of a , we can define the integral on $(a, b]$ in the same way. If f is not bounded in any neighborhood of a or b and if there exists a point c ($a < c < b$) for which the improper integrals $\int_a^c f(x) dx$ and $\int_c^b f(x) dx$ exist, then we define $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$, which is independent of the choice of the point c . Fur-

thermore, assume that $f(x)$ is not bounded in any neighborhood of each point c_j ($j = 1, \dots, n$) ($a < c_1 < \dots < c_n < b$). Then we define

$$\begin{aligned} \int_a^b f(x) dx = \int_a^{c_1} f(x) dx + \int_{c_1}^{c_2} f(x) dx + \dots \\ + \int_{c_{n-1}}^{c_n} f(x) dx + \int_{c_n}^b f(x) dx, \end{aligned}$$

provided that all improper integrals

$$\int_a^{c_1} f(x) dx, \dots, \int_{c_n}^b f(x) dx$$

exist. Suppose that $f(x)$ is defined on $[a, b]$ and bounded outside any neighborhood of $c \in (a, b)$ but not bounded in either $[c - \varepsilon, c]$ or $[c, c + \varepsilon]$ for any $\varepsilon > 0$. It may well happen that, although neither $\lim_{\varepsilon \rightarrow 0} \int_a^{c-\varepsilon} f(x) dx$ nor $\lim_{\varepsilon' \rightarrow 0} \int_{c+\varepsilon'}^b f(x) dx$ exists (accordingly, the improper integral $\int_a^b f(x) dx$ does not exist), if we put $\varepsilon = \varepsilon'$, the limit

$$\lim_{\varepsilon \rightarrow 0} \left(\int_a^{c-\varepsilon} f(x) dx + \int_{c+\varepsilon}^b f(x) dx \right)$$

does exist. This limit is called **Cauchy's principal value** and is denoted by p.v. $\int_a^b f(x) dx$ (v.p. in French). For example, p.v. $\int_{-1}^1 (dx/x) = \lim_{\varepsilon \rightarrow 0} (\int_{-1}^{-\varepsilon} (1/x) dx + \int_{\varepsilon}^1 (1/x) dx) = 0$.

E. Integrals on Infinite Intervals

Suppose that we are given a function $f(x)$ defined on an infinite interval $[a, \infty)$ and integrable on any finite interval $[a, b]$. If $\lim_{b \rightarrow \infty} \int_a^b f(x) dx$ exists and is finite, then this limit is called the **improper integral** of f on $[a, \infty)$ and is denoted by $\int_a^\infty f(x) dx$. We define similarly $\int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$, where f is defined on $(-\infty, b]$ and integrable on any interval $[a, b]$. Furthermore, $\int_{-\infty}^\infty f(x) dx$ is, by definition, $\int_{-\infty}^c f(x) dx + \int_c^\infty f(x) dx$, which is independent of the choice of c . Suppose that $f(x)$ is integrable on $[a, b]$ for a fixed a and an arbitrary b larger than a . If $f(x) = O(x^\alpha)$ for some $\alpha < -1$, then $\int_a^\infty f(x) dx$ exists. Generally, for α, β such that $-\infty \leq \alpha < \infty$ and $-\infty < \beta \leq \infty$, if the improper integral $\int_\alpha^\beta f(x) dx$ exists, we say that the integral is **convergent**; otherwise, it is **divergent**. Improper integrals also satisfy the three basic properties of integrals (1), (2), and (3) (\rightarrow Section B). However, the existence of an improper integral of a function f on an interval I does not imply the existence of the improper integral of $|f|$ on the same interval I . For example, let f be a function determined by $f(0) = 0$, $f(x) = (1/x) \sin(1/x)$ for $0 < x \leq \pi$. Then $\int_0^\pi f(x) dx$ exists, but $\int_0^\pi |f(x)| dx$ does not. On the other hand, if the improper integral of $|f(x)|$ exists,

then the improper integral of $f(x)$ exists, and we have

$$\left| \int_{\alpha}^{\beta} f(x) dx \right| \leq \int_{\alpha}^{\beta} |f(x)| dx,$$

where $-\infty \leq \alpha < \beta \leq +\infty$. In this case, we say that f is **absolutely integrable** on the interval $[\alpha, \beta]$. Assume now that $f(x)$ is defined on $(-\infty, \infty)$ and integrable on any finite interval. If $\lim_{x \rightarrow \infty} \int_{-x}^x f(x) dx$ exists, then it is called **Cauchy's principal value** of the integral of f in $(-\infty, \infty)$.

If $f(x)$ is a monotone decreasing, positive, and continuous function defined on $[k, \infty)$ (where k is an integer), then according as $\sum_{v=k}^{\infty} f(v)$ converges or diverges, so does $\int_k^{\infty} f(x) dx$.

Suppose that a series $\sum_{n=1}^{\infty} f_n(x)$, where all the functions $f_n(x)$ ($n = 1, 2, \dots$) are defined and nonnegative on an infinite interval $[a, \infty)$, satisfies $\int_a^b (\sum_{n=1}^{\infty} f_n(x)) dx = \sum_{n=1}^{\infty} \int_a^b f_n(x) dx$ for arbitrary $b > a$. Then according as $\sum_{n=1}^{\infty} \int_a^{\infty} f_n(x) dx$ converges or diverges, so does $\int_a^{\infty} (\sum_{n=1}^{\infty} f_n(x)) dx$. When they converge, the following equality holds: $\int_a^{\infty} \sum_{n=1}^{\infty} f_n(x) dx = \sum_{n=1}^{\infty} \int_a^{\infty} f_n(x) dx$. In this theorem, if $\int_a^{\infty} \sum_{n=1}^{\infty} |f_n(x)| dx$ or $\sum_{n=1}^{\infty} \int_a^{\infty} |f_n(x)| dx$ converges, then the same conclusion as above will follow even when the $f_n(x)$ are not necessarily positive (\rightarrow Appendix A, Table 9).

F. Multiple Integrals

Suppose that $f(x, y)$ is a function defined and bounded on an interval $I = \{(x, y) | a \leq x \leq b, c \leq y \leq d\}$ in the xy -plane. Partitions $\{x_j\}$ and $\{y_k\}$ of $[a, b]$ and $[c, d]$ with $a = x_0 < x_1 < \dots < x_m = b$ and $c = y_0 < y_1 < \dots < y_n = d$ determine a "partition," denoted by D , of I into subintervals of the form $I_{jk} = \{(x, y) | x_{j-1} \leq x \leq x_j, y_{k-1} \leq y \leq y_k\}$ ($j = 1, \dots, m; k = 1, \dots, n$). Writing

$$M_{jk} = \sup_{(x, y) \in I_{jk}} f(x, y),$$

$$m_{jk} = \inf_{(x, y) \in I_{jk}} f(x, y),$$

we set

$$\bar{\sigma}(D) = \sum_{j=1}^m \sum_{k=1}^n M_{jk} (x_j - x_{j-1}) (y_k - y_{k-1}),$$

$$\underline{\sigma}(D) = \sum_{j=1}^m \sum_{k=1}^n m_{jk} (x_j - x_{j-1}) (y_k - y_{k-1}).$$

Then we obtain $\inf_D \bar{\sigma}(D) \geq \sup_D \underline{\sigma}(D)$. If $\inf_D \bar{\sigma}(D) = \sup_D \underline{\sigma}(D)$, then $f(x, y)$ is called integrable on I , and the common value is called the **double integral** of f on I and is denoted by $\iint_I f(x, y) dx dy$. Analogously, we can define **n -tuple integrals** (or **multiple integrals**) and the integrability of functions of n variables.

Let K be a bounded set in the xy -plane and

I an interval containing K . Let $\varphi(x, y)$ be the characteristic function of K defined on I , that is, φ is determined by

$$\varphi(x, y) = 1 \quad \text{for } (x, y) \in K,$$

$$\varphi(x, y) = 0 \quad \text{for } (x, y) \in I - K.$$

Replacing $f(x, y)$ by this $\varphi(x, y)$, we consider $\inf_D \bar{\sigma}(D)$ ($\sup_D \underline{\sigma}(D)$). These values can be shown to be independent of the choice of such an interval I and are called the **outer area** (**inner area**) of K , respectively. When these two values coincide, K is said to be of **definite area**, and the common value is called the **area of K** . A necessary and sufficient condition for K to be of definite area is that the outer area of the \dagger boundary of K be zero. Now consider a bounded function defined on a set K of definite area. Then, taking an interval I containing K , define an extension $\varphi(x, y)$ of $f(x, y)$ as follows:

$$\varphi(x, y) = f(x, y) \quad \text{for } (x, y) \in K,$$

$$\varphi(x, y) = 0 \quad \text{for } (x, y) \in I - K.$$

If $\varphi(x, y)$ is integrable on I , then $f(x, y)$ is called **integrable** on K , and the integral of f on K is defined by $\iint_K f(x, y) dx dy = \iint_I \varphi(x, y) dx dy$, which is independent of the special choice of I . The set K is called the **domain of integration**. Since K is of definite area, the set of boundary points of K at which $\varphi(x, y)$ is not continuous can be contained in a union of intervals whose total area can be made smaller than any preassigned positive number. Consequently, a function bounded on K and continuous at each \dagger interior point of K is integrable on K . Like integrals of functions of a single variable, multiple integrals satisfy the three basic properties of integrals (\rightarrow Section B).

G. Multiple Integrals and Iterated Integrals

Suppose that we are given a function $f(x, y)$ that is continuous on an interval $I = \{(x, y) | a \leq x \leq b, c \leq y \leq d\}$. Then, for a fixed y in $[c, d]$, the function $f(x, y)$, regarded as a function of x , can be integrated with respect to x on the interval $[a, b]$, and the integral thus obtained is a continuous function of y . The integral of the function defined on $[c, d]$, namely, $\int_c^d (\int_a^b f(x, y) dx) dy$, is called the **iterated integral** (or **repeated integral**) of $f(x, y)$ and is often written as $\int_c^d dy \int_a^b f(x, y) dx$. The following formula gives a representation of a double integral by iterated ones:

$$\begin{aligned} \iint_I f(x, y) dx dy &= \int_c^d dy \int_a^b f(x, y) dx \\ &= \int_a^b dx \int_c^d f(x, y) dy. \end{aligned}$$

More generally, $\varphi_1(x)$ and $\varphi_2(x)$ being continuous on $[a, b]$ and $\varphi_1(x) \leq \varphi_2(x)$, consider the following subset $K = \{(x, y) | a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}$ of the xy -plane. Suppose further that $f(x, y)$ is continuous on K . Then the following equality holds:

$$\iint_K f(x, y) dx dy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy.$$

In the case of unbounded integrands or unbounded domains of integration, we can still define integrals under suitable restrictions. For instance, assume the following two properties: (1) There exists a sequence $\{K_n\}$ of sets, each of which is of definite area, satisfying $K_1 \subset K_2 \subset \dots$ and $K = \bigcup_{n=1}^{\infty} K_n$. (2) $f(x, y)$ is bounded and integrable on each K_n ($n = 1, 2, \dots$). If a finite limit $\lim_{n \rightarrow \infty} \iint_{K_n} f(x, y) dx dy$ exists and is independent of the choice of $\{K_n\}$, then $f(x, y)$ is called integrable on K . This limit is called the integral of $f(x, y)$ on K and is denoted by $\iint_K f(x, y) dx dy$:

$$\lim_{n \rightarrow \infty} \iint_{K_n} f(x, y) dx dy = \iint_K f(x, y) dx dy.$$

When the integral thus defined exists, we say that the integral is convergent. If a finite limit $\lim_{n \rightarrow \infty} \iint_{K_n} |f(x, y)| dx dy$ exists for some sequence $\{K_n\}$ with property (1), then f is integrable on K . Let $f(x, y)$ be continuous and nonnegative on $K = \{(x, y) | \alpha < x < \beta, \gamma < y < \delta\}$, where $-\infty \leq \alpha < \beta \leq \infty, -\infty \leq \gamma < \delta \leq \infty$. Furthermore, let $f(x, y)$ be integrable on K , and assume that the improper integral $F(x) = \int_{\gamma}^{\delta} f(x, y) dy = \lim_{c \downarrow \gamma, d \uparrow \delta} \int_c^d f(x, y) dy$ exists and converges uniformly with respect to x as $c \downarrow \gamma, d \uparrow \delta$. Then $\int_{\alpha}^{\beta} F(x) dx$ is well defined, and we have

$$\iint_K f(x, y) dx dy = \int_{\alpha}^{\beta} dx \int_{\gamma}^{\delta} f(x, y) dy.$$

In particular, if $\alpha = a, \gamma = b, \beta = \delta = \infty$, then we have

$$\int_a^{\infty} \int_b^{\infty} f(x, y) dx dy = \int_a^{\infty} dx \int_b^{\infty} f(x, y) dy.$$

H. Interchanging the Order of Differentiation and Integration

If both $f(x, y)$ and $\partial f(x, y)/\partial y$ are continuous on an interval $\{(x, y) | a \leq x \leq b, y_0 - \eta \leq y \leq y_0 + \eta\}$, then we can interchange the order of differentiation and integration as follows:

$$\frac{d}{dy} \int_a^b f(x, y) dx = \int_a^b \frac{\partial f(x, y)}{\partial y} dx \quad \text{for } y = y_0.$$

Assume further that this equality holds for

216 Ref. Integral Calculus

every $b(>a)$, the improper integral

$$\int_a^{\infty} f(x, y) dx = \lim_{b \rightarrow \infty} \int_a^b f(x, y) dx$$

converges, and the improper integral

$$\int_a^{\infty} \frac{\partial f(x, y)}{\partial y} dx = \lim_{b \rightarrow \infty} \int_a^b \frac{\partial f(x, y)}{\partial y} dx$$

converges as $b \rightarrow \infty$ uniformly for y with $|y - y_0| < \eta$. Then

$$\frac{d}{dy} \int_a^{\infty} f(x, y) dx = \int_a^{\infty} \frac{\partial f(x, y)}{\partial y} dx \quad \text{for } y = y_0.$$

Several other similar theorems are known.

Though the previous theorems are written in terms of two variables, analogous theorems hold for n variables.

I. Change of Variables in Multiple Integrals

Let G be a bounded domain of definite area in an n -dimensional Euclidean space $\mathbf{R}^n(x)$. Assume that a mapping $x \rightarrow y(x) = (y_1(x_1, \dots, x_n), \dots, y_n(x_1, \dots, x_n))$ is of class C^1 from an open set containing the closure \bar{G} of G into an n -dimensional Euclidean space $\mathbf{R}^n(y)$. We denote the image of G under this mapping by B . If $f(y_1, \dots, y_n)$ is continuous on B , then the following formula on change of variables holds:

$$\int \dots \int_B f(y_1, \dots, y_n) dy_1 \dots dy_n = \int \dots \int_G g(x_1, \dots, x_n) \left| \frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)} \right| dx_1 \dots dx_n,$$

where $g(x_1, \dots, x_n) = f(y_1(x_1, \dots, x_n), \dots, y_n(x_1, \dots, x_n))$ and $D(y_1, \dots, y_n)/D(x_1, \dots, x_n)$ is the

*Jacobian determinant of the mapping $y(x)$. This formula is usually utilized in the case where y_1, \dots, y_n are *functionally independent, though otherwise both sides vanish and the formula still holds. For improper integrals, a similar formula will hold under suitable restrictions, for example, if the integrals converge absolutely.

For related topics \rightarrow 94 Curvilinear Integrals and Surface Integrals, 221 Integration Theory, and 270 Measure Theory.

References

- [1] T. M. Apostol, *Mathematical analysis*, Addison-Wesley, 1957.
- [2] N. Bourbaki, *Eléments de mathématique, Fonctions d'une variable réelle*, Actualités Sci. Ind., 1074b, 1132a, Hermann, second edition, 1958, 1961.
- [3] R. C. Buck, *Advanced calculus*, McGraw-Hill, second edition, 1965.

- [4] R. Courant, Differential and integral calculus I, II, Nordemann, 1938.
 [5] G. H. Hardy, A course of pure mathematics, Cambridge Univ. Press, seventh edition, 1938.
 [6] E. Hille, Analysis I, II, Blaisdell, 1964, 1966.
 [7] W. Kaplan, Advanced calculus, Addison-Wesley, 1952.
 [8] E. Landau, Einführung in die Differentialrechnung und Integralrechnung, Noordhoff, 1934; English translation, Differential and integral calculus, Chelsea, 1965.
 [9] J. M. H. Olmsted, Advanced calculus, Appleton-Century-Crofts, 1961.
 [10] A. Ostrowski, Vorlesungen über Differential- und Integralrechnung I–III, Birkhäuser, second edition, 1960–1961.
 [11] M. H. Protter and C. B. Morrey, Modern mathematical analysis, Addison-Wesley, 1964.
 [12] W. Rudin, Principles of mathematical analysis, McGraw-Hill, second edition, 1964.
 [13] V. I. Smirnov, A course of higher mathematics. I, Elementary calculus; II, Advanced calculus, Addison-Wesley, 1964.
 [14] A. E. Taylor, Advanced calculus, Ginn, 1955.

217 (XIII.30) Integral Equations

A. General Remarks

Equations including the integrals of unknown functions are called **integral equations**. The most studied ones are the **linear integral equations**, i.e., linear in unknown functions.

Let D be a domain of n -dimensional Euclidean space and $f(x)$ and $K(x, y)$ be functions defined for $x = (x_1, x_2, \dots, x_n) \in D$, $y = (y_1, y_2, \dots, y_n) \in D$. **Integral equations of Fredholm type** (or **Fredholm integral equations**) [1] are those of the forms

$$\int_D K(x, y)\varphi(y)dy = f(x), \quad (1)$$

$$\varphi(x) - \int_D K(x, y)\varphi(y)dy = f(x), \quad (2)$$

$$A(x)\varphi(x) - \int_D K(x, y)\varphi(y)dy = f(x), \quad (3)$$

where $\varphi(x)$ is an unknown function and $\int_D dy$ means the n -fold integral $\int \dots \int_D dy_1 \dots dy_n$. Equations of the forms (1), (2), and (3) are called equations of the **first**, **second**, and **third kind**, respectively. Equations of the second kind have been investigated in great detail.

Equations of the third kind, in many cases, can be reduced formally to those of the second kind. The function $K(x, y)$ is called a **kernel** (or **integral kernel**) of the integral equation.

Integral equations of Volterra type (or **Volterra integral equations**) are those of the forms

$$\int_a^x K(x, y)\varphi(y)dy = f(x), \quad (1')$$

$$\varphi(x) - \int_a^x K(x, y)\varphi(y)dy = f(x), \quad (2')$$

$$A(x)\varphi(x) - \int_a^x K(x, y)\varphi(y)dy = f(x), \quad (3')$$

where $\varphi(x)$ is an unknown function. Equations of the forms (1'), (2'), and (3') are also called equations of the **first**, **second**, and **third kind**, respectively. Integral equations of Volterra type can be regarded as integral equations of Fredholm type having kernels equal to 0 for $x < y$, but these two types of equations are usually treated separately, since they have considerably different characters.

The kernels in equations (1)–(3) and (1')–(3') are frequently written in the form $\lambda K(x, y)$ with a parameter λ , in particular when the equations are related to eigenvalue problems, which is explained in Section F.

The theory of integral equations was originated in 1823 by N. H. Abel, who investigated the relationship between time and the path of a falling body in the field of gravitation. Let $\varphi(t)$ be a quantity varying with time, which is connected by some law with its values in some time interval of the past or the future. Then the law of variation of $\varphi(t)$ can be described mathematically by an integral equation. The situation is the same even when the variable t is not time but a coordinate of the space. In this way, various problems in physics can be reduced to solutions of integral equations.

B. Relation to Differential Equations

Many problems in differential equations can be reduced to problems related to integral equations. Such reduction often makes the problems easier to handle and clarifies the nature of the solutions. For example, consider the problem of finding a solution of the ordinary second-order linear differential equation $d^2y/dx^2 + \lambda y = 0$ with the boundary condition $y(0) = y(1) = 0$ [4]. Let $d^2y/dx^2 = u(x)$. If we integrate the equation twice, change the order of integration, and make use of the boundary condition, then we have

$$y = \int_0^x (x - \xi)u(\xi)d\xi - \int_0^1 x(1 - \xi)u(\xi)d\xi,$$

from which we see that the given differential equation can be written in the form

$$u = \lambda \int_0^1 x(1-\xi)u(\xi)d\xi - \lambda \int_0^x (x-\xi)u(\xi)d\xi.$$

Decomposing the first integral on the right-hand side into the sum of an integral over $(0, x)$ and one over $(x, 1)$, and combining the integral over $(0, x)$ with the second integral on the right-hand side, we obtain a Fredholm integral equation of the first kind as follows:

$$u = \lambda \int_0^1 G(x, \xi)u(\xi)d\xi,$$

$$G(x, \xi) = \begin{cases} \xi(1-x) & (0 \leq \xi \leq x), \\ x(1-\xi) & (x \leq \xi \leq 1). \end{cases}$$

Clearly, the solution of this integral equation is equivalent to that of the original differential equation. The function G is called †Green's function for the boundary value condition $y(0) = y(1) = 0$ in the theory of boundary value problems. Differential equations of higher orders can be treated analogously (→ 315 Ordinary Differential Equations (Boundary Value Problems)). †Initial value problems of linear ordinary differential equations can be reduced to the solution of Volterra integral equations in a similar way.

As another example [5, 6], consider the †Dirichlet problem on a plane, i.e., the problem of finding a function u satisfying the conditions (i) u is †harmonic in the interior of the region D bounded by a closed curve C ($\xi = \varphi(s)$, $\eta = \psi(s)$, $0 \leq s \leq l$); (ii) $u(x, y) \rightarrow F(s)$ uniformly with respect to (x_0, y_0) as (x, y) approaches (x_0, y_0) from the inside of D , where (x_0, y_0) is an arbitrary point on C , $F(s)$ is a continuous function given on C , and s is the arc length along C . Put $f(s) = F(s)/\pi$ and

$$K(s; t) = \frac{(\psi(s) - \psi(t))\varphi'(t) - (\varphi(s) - \varphi(t))\psi'(t)}{\pi((\varphi(s) - \varphi(t))^2 + (\psi(s) - \psi(t))^2)}.$$

Then it is known that a solution u of the problem can be given in the form

$$u(x, y) = \int_0^l \mu(s) \frac{\partial}{\partial n} \log \frac{1}{r} ds,$$

where $r^2 = (\varphi(s) - x)^2 + (\psi(s) - y)^2$, n is the inner normal of C , and $\mu(s)$ is a continuous solution of the following Fredholm integral equation of the second kind:

$$\mu(s) = f(s) - \int_0^l K(s; t)\mu(t)dt.$$

We can treat the †Neumann problem similarly, i.e., the problem in which condition (ii) is replaced by (ii') $(\partial u / \partial n)(x, y) \rightarrow F(s)$ uniformly with respect to (x_0, y_0) as (x, y) approaches (x_0, y_0) from the inside of D . In the Neumann

problem, put $f(s) = F(s)/\pi$ and

$$L(s; t) = \frac{(\psi(s) - \psi(t))\varphi'(s) - (\varphi(s) - \varphi(t))\psi'(s)}{\pi((\varphi(s) - \varphi(t))^2 + (\psi(s) - \psi(t))^2)}.$$

Then we have a solution u in the form

$$u(x, y) = - \int_0^l \mu(s) \log \frac{1}{r} ds + c,$$

where $\mu(s)$ is a solution of the following Fredholm integral equation of the second kind:

$$\mu(s) = f(s) - \int_0^l L(s; t)\mu(t)dt.$$

A solution of this integral equation, however, exists when and only when $\int_0^l F(s)ds = 0$. In the expression of a solution $u(x, y)$ of the Neumann problem, c is an arbitrary additive constant, up to which a solution of the problem is determined uniquely. We can also treat partial differential equations of †elliptic type in an analogous way.

C. Integral Equations with Continuous Kernel

We describe some results for integral equations with m -dimensional independent variables, i.e., equations in which D is an m -dimensional closed domain. We assume that $K(x, y)$ and $f(x)$ are continuous in Sections D–H.

D. The Method of Successive Iteration

Among methods of solving Fredholm integral equations of the second kind, the simplest is the **method of successive iteration**, sometimes called the **method of successive approximation** [7]. In the method of successive iteration, we rewrite (2) in the form

$$\varphi(x) = f(x) + \int_D K(x, y)\varphi(y)dy$$

and replace the function $\varphi(y)$ on the right-hand side by the function

$$f(y) + \int_D K(y, z)\varphi(z)dz.$$

If we repeat the process successively, then we have

$$\varphi(x) = f(x) + \sum_{i=1}^n \int_D K_i(x, y)f(y)dy + \int_D K_{n+1}(x, y)\varphi(y)dy,$$

where

$$K_1(x, y) = K(x, y),$$

$$K_i(x, y) = \int_D K_{i-1}(x, s)K(s, y)ds.$$

The functions $K_i(x, y)$ are called the **iterated kernels**. Assume that $\sum_{n=1}^{\infty} K_n(x, y)$ converges uniformly. Then, putting

$$R(x, y) = \sum_{n=1}^{\infty} K_n(x, y), \quad (4)$$

we obtain a solution of (2) in the form

$$\varphi(x) = f(x) + \int_D R(x, y) f(y) dy. \quad (5)$$

The series (4) is called a **Neumann series**.

For a given kernel $K(x, y)$, a function $R(x, y)$ satisfying

$$K(x, y) - R(x, y) + \int_D K(x, s) R(s, y) ds = 0$$

and

$$K(x, y) - R(x, y) + \int_D R(x, s) K(s, y) ds = 0$$

is called a **resolvent** of $K(x, y)$ (in some cases $-R(x, y)$ is taken as the resolvent). If a resolvent of $K(x, y)$ exists, the solution of (2) can be given uniquely by (5). If a Neumann series converges uniformly, (4) gives a resolvent of $K(x, y)$.

If we apply a similar process to Volterra integral equations of the second kind, then we have the iterated kernels defined by

$$K_{i+1}(x, y) = \int_y^x K_i(x, s) K(s, y) ds \quad (i = 1, 2, \dots).$$

For these iterated kernels, a Neumann series defined by (4) always converges uniformly.

E. Fredholm's Method

Let D be a bounded closed domain and $K(x, y)$ a continuous kernel. A Neumann series (4) converges uniformly and gives a resolvent if $|K(x, y)|$ or the region D is sufficiently small, but otherwise it does not necessarily converge.

E. I. Fredholm [1, 7] gave a method of constructing a resolvent for the more general case. Write a kernel in the form $\lambda K(x, y)$, and put

$$K \begin{pmatrix} x_1, \dots, x_n \\ y_1, \dots, y_n \end{pmatrix} = \begin{vmatrix} K(x_1, y_1) & \dots & K(x_1, y_n) \\ \vdots & \ddots & \vdots \\ K(x_n, y_1) & \dots & K(x_n, y_n) \end{vmatrix}.$$

Define $D(\lambda)$ and $D(x, y; \lambda)$ by

$$D(\lambda) = 1 + \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \int_D \dots \int_D K \begin{pmatrix} s_1, \dots, s_n \\ s_1, \dots, s_n \end{pmatrix} ds_1 \dots ds_n$$

and

$$D(x, y; \lambda) = K(x, y) + \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \int_D \dots \int_D K \begin{pmatrix} x, s_1, \dots, s_n \\ y, s_1, \dots, s_n \end{pmatrix} ds_1 \dots ds_n.$$

The series in these two equations both converge uniformly and hence define entire functions of λ . The functions $D(\lambda)$ and $D(x, y; \lambda)$ are called **Fredholm's determinant** and **Fredholm's first minor** of the kernel $K(x, y)$, respectively.

For small $|\lambda|$, we have

$$\frac{D'(\lambda)}{D(\lambda)} = \sum_{n=1}^{\infty} \lambda^{n-1} \int_D K_n(s, s) ds,$$

where the $K_n(x, y)$ are iterated kernels corresponding to $K(x, y)$. Now if $D(\lambda) \neq 0$, a resolvent $\lambda R(x, y; \lambda)$ of the kernel $\lambda K(x, y)$ can be given by

$$\frac{D(x, y; \lambda)}{D(\lambda)} = R(x, y; \lambda).$$

If $D(\lambda) = 0$, some extension of the method in this section is needed. Fredholm introduced for this purpose **Fredholm's r th minor**

$$D \begin{pmatrix} x_1, \dots, x_r, \\ y_1, \dots, y_r \end{pmatrix}; \lambda$$

defined by

$$\begin{aligned} & D \begin{pmatrix} x_1, \dots, x_r, \\ y_1, \dots, y_r \end{pmatrix}; \lambda \\ &= K \begin{pmatrix} x_1, \dots, x_r \\ y_1, \dots, y_r \end{pmatrix} + \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \\ & \quad \times \int_D \dots \int_D K \begin{pmatrix} x_1, \dots, x_r, s_1, \dots, s_n \\ y_1, \dots, y_r, s_1, \dots, s_n \end{pmatrix} ds_1 \dots ds_n. \end{aligned}$$

F. Eigenvalue Problems and Fredholm's Alternative Theorem

Consider a **homogeneous integral equation** of the second kind

$$\varphi(x) - \lambda \int_D K(x, y) \varphi(y) dy = 0, \quad (6)$$

where D is a bounded closed domain and $K(x, y)$ is continuous in $D \times D$. When (6) has a nontrivial solution $\varphi(x)$ for some λ , then λ is called an **eigenvalue** corresponding to the kernel $K(x, y)$, and the corresponding nontrivial solution $\varphi(x)$ is called an **eigenfunction** corresponding to the kernel $K(x, y)$. If $D(\lambda) \neq 0$, then (6) has no nontrivial solution, from which it follows that eigenvalues must be zero points of the entire function $D(\lambda)$. For an arbitrary eigenvalue λ , there is a set of linearly independent eigenfunctions corresponding to λ such that any eigenfunction corresponding to λ can be written as a linear combination of the eigenfunctions belonging to the set under consideration. Such a set of linearly independent eigenfunctions corresponding to an eigenvalue λ is called a **fundamental system** corresponding to the eigenvalue λ . The number of elements of the fundamental system is called

the **index of the eigenvalue** λ . The index of an eigenvalue is always finite. The homogeneous integral equation of the form

$$\psi(y) - \lambda \int_D K(x, y) \psi(x) dx = 0 \quad (6')$$

is called an **associated (or transposed) integral equation** of (6). The associated equation has the same eigenvalues as the original equation; moreover, the index of a common eigenvalue is the same for both equations. For any eigenvalue λ , the order of the zero point λ of the entire function $D(\lambda)$ is called the **multiplicity of the eigenvalue** λ . If an eigenvalue λ is a pole of $R(x, y; \lambda)$, then we have $p + 1 \geq r + q$, where r is the order of the pole, p is the multiplicity of λ , and q is the index of λ . In particular, if λ is a simple pole of $R(x, y; \lambda)$, we have $p = q$, namely, the multiplicity is equal to the index. An example with this particular property is the integral equation with a symmetric kernel to be discussed in Section G. For the set of eigenvalues there is no finite accumulation point even if there are infinitely many eigenvalues.

If λ is not an eigenvalue, the inhomogeneous equation

$$\varphi(x) - \lambda \int_D K(x, y) \varphi(y) dy = f(x) \quad (7)$$

can be solved uniquely for any continuous function $f(x)$. In this case we have $D(\lambda) \neq 0$, and the resolvent $R(x, y; \lambda)$ of the kernel $\lambda K(x, y)$ exists. If λ is an eigenvalue, we have $D(\lambda) = 0$, and equation (7) has a solution if and only if

$$\int_D \psi(x) f(x) dx = 0$$

for all solutions $\psi(y)$ of (6') (linearly independent solutions $\psi(y)$ are finite in number). The last statement is called **Fredholm's alternative theorem** (— 68 Compact and Nuclear Operators).

A kernel of the type

$$K(x, y) = \sum_{j=1}^n X_j(x) Y_j(y)$$

is called a **separated kernel**, **degenerate kernel**, or **Pincherle-Goursat kernel**. For such a kernel, we have $D(\lambda) = \det(\delta_{jk} - \lambda \int_a^b X_j(t) Y_k(t) dt)$, and hence we can easily obtain eigenvalues and eigenfunctions. A nondegenerate kernel can be studied using the results obtained for separated kernels, since we can regard a kernel of the general form as the limit of a sequence of separated kernels.

G. Symmetric Kernels

A kernel $K(x, y)$ is called a **symmetric kernel** if it is real and $K(x, y) = K(y, x)$. Let D be a

bounded closed domain and $K(x, y)$ be a continuous symmetric kernel. In this case the associated equation (6') clearly coincides with the original equation (6) [5, 6, 8].

Corresponding to any nontrivial symmetric kernel $K(x, y)$, there exist at least one eigenvalue and one eigenfunction. The eigenvalues are all real, and the eigenfunctions corresponding to distinct eigenvalues are mutually orthogonal. If we orthonormalize the eigenfunctions belonging to all fundamental systems and number them according to the order of increasing absolute values of the corresponding eigenvalues, then we have an orthonormal system $\{\varphi_i(x)\}$, called a **complete orthonormal system of fundamental functions** or simply a **complete orthogonal (or orthonormal) system**. If we number the eigenvalues taking their multiplicities into account and according to the order of increasing absolute values, then we have the equality

$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i^2} = \int_D \int_D K^2(x, y) dx dy.$$

Corresponding to an iterated kernel $K_n(x, y)$, we have the eigenvalues $\{\lambda_i^n\}$, and we can choose the corresponding orthonormal system so that it coincides with the one corresponding to $K(x, y)$. Eigenvalues and eigenfunctions corresponding to an iterated kernel can be obtained in the following way: Put $\int_D K_n(s, s) ds = u_n$; then the following limit exists:

$$\lim_{n \rightarrow \infty} u_{2n}/u_{2n+2} = \lambda^2 < +\infty,$$

which gives an eigenvalue of the iterated kernel $K_2(x, y)$. A function $\varphi(x, y)$ defined by

$$\lim_{n \rightarrow \infty} \lambda^{2n} K_{2n}(x, y) = \varphi(x, y)$$

(uniformly convergent) gives the corresponding eigenfunction $\varphi(x, c)$ for any constant c satisfying $\varphi(x, c) \neq 0$.

Let λ^n be an eigenvalue corresponding to an iterated kernel $K_n(x, y)$ and $\varphi(x)$ be a corresponding eigenfunction. Consider the functions $\psi_j(x)$ ($j = 0, 1, \dots, n-1$) defined by

$$\psi_j(x) = \varphi(x) + \sum_{k=1}^{n-1} \varepsilon^{kj} \lambda^k \int_D K_k(x, y) \varphi(y) dy \quad (j = 0, 1, 2, \dots, n-1),$$

where ε is one of the n th primitive roots of 1. Then for at least one value of j , $\varepsilon^j \lambda$ is an eigenvalue corresponding to the kernel $K(x, y)$, and ψ_j is a corresponding eigenfunction. This relationship between eigenvalues and eigenfunctions corresponding to an iterated kernel and those corresponding to the original kernel is valid even for kernels that are not necessarily symmetric [2].

Let a kernel $K^{(n)}(x, y)$ be defined by

$$K^{(n)}(x, y) = K(x, y) - \sum_{i=1}^n \frac{\varphi_i(x)\varphi_i(y)}{\lambda_i},$$

where the λ_i ($i = 1, 2, \dots, n$) are the eigenvalues corresponding to the kernel $K(x, y)$ and the $\varphi_i(x)$ ($i = 1, \dots, n$) are the corresponding orthonormalized eigenfunctions. Then eigenvalues and eigenfunctions corresponding to $K^{(n)}(x, y)$ are those corresponding to $K(x, y)$, with the exception of $\lambda_1, \dots, \lambda_n$ and $\varphi_1(x), \dots, \varphi_n(x)$.

Let $\varphi(x)$ be any function that satisfies $\int_D (\varphi(x))^2 dx = 1$. Then the integral

$$J = \int_D \int_D K_2(x, y) \varphi(x) \varphi(y) dx dy$$

assumes the maximum value when $\varphi(x)$ is an eigenfunction corresponding to $K_2(x, y)$ with the smallest eigenvalue λ_1^2 . Let the eigenvalues λ_n of $K(x, y)$ be numbered in order of increasing absolute values, so that $|\lambda_n| \leq |\lambda_{n+1}|$. Let $\varphi(x)$ be any function satisfying

$$\int_D \psi_i(x) \varphi(x) dx = 0 \quad (i = 1, 2, \dots, n),$$

$$\int_D (\varphi(x))^2 dx = 1$$

for given functions $\psi_i(x)$ ($i = 1, \dots, n$). Then the maximum value of the integral J above is the least when the set of all linear combinations of $\{\psi_1(x), \dots, \psi_n(x)\}$ coincides with the set of all linear combinations of $\{\varphi_1(x), \dots, \varphi_n(x)\}$, and in this case the maximum of J is attained by some eigenfunction $\varphi(x)$ corresponding to $K_2(x, y)$ with the eigenvalue λ_{n+1}^2 . The results in this paragraph show that we can obtain eigenvalues by solving a variational problem concerning the integral J .

H. Expansion Theorems

Let $K(x, y)$ be a continuous symmetric kernel and $h(x)$ be a function square integrable on a bounded closed domain D . Then a function $f(x)$ such that

$$f(x) = \int_D K(x, y) h(y) dy$$

can be expanded in the form

$$f(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x),$$

where $\{\varphi_i(x)\}$ is a complete orthonormal system of fundamental functions corresponding to $K(x, y)$ and

$$c_n = \int_D f(x) \varphi_n(x) dx \quad (n = 1, 2, \dots).$$

The series in the expansion of $f(x)$ converges uniformly. These facts are the content of the **Hilbert-Schmidt expansion theorem** [5, 6, 8]. By using this theorem, for a λ that is not an eigenvalue, we can obtain a solution $\varphi(x)$ of the Fredholm integral equation (7) with a symmetric kernel in the form

$$\varphi(x) = f(x) + \lambda \sum_{i=1}^{\infty} \frac{\varphi_i(x)}{\lambda_i - \lambda} \int_D f(x) \varphi_i(x) dx.$$

For $m \geq 2$, iterated kernels can be expanded in the form

$$K_m(x, y) = \sum_{i=1}^{\infty} \frac{\varphi_i(x) \varphi_i(y)}{\lambda_i^m}$$

(uniformly convergent). If $\lambda R(x, y; \lambda)$ is a resolvent of a symmetric kernel $\lambda K(x, y)$, then $R(x, y; \lambda)$ can be expanded as

$$R(x, y; \lambda) = K(x, y) + \lambda \sum_{i=1}^{\infty} \frac{\varphi_i(x) \varphi_i(y)}{\lambda_i(\lambda_i - \lambda)}.$$

If a symmetric kernel $K(x, y)$ satisfies the inequality

$$\int_D \int_D K(x, y) \varphi(x) \varphi(y) dx dy \geq 0$$

for all $\varphi(x)$, it is called a **positive (semidefinite) kernel**. If in this inequality the equality holds only for $\varphi(x) \equiv 0$, $K(x, y)$ is called a **positive definite kernel**. For a positive definite kernel, eigenvalues are all positive, and the kernel can be expanded in the form

$$K(x, y) = \sum_{i=1}^{\infty} \frac{\varphi_i(x) \varphi_i(y)}{\lambda_i}$$

(uniformly convergent). This result is called **Mercer's theorem**.

When a real continuous kernel $K(x, y)$ is not symmetric, we consider two positive kernels $\hat{K}'(x, y)$ and $\hat{K}''(x, y)$ defined by

$$\int_D K(x, s) K(y, s) ds = \hat{K}'(x, y)$$

and

$$\int_D K(s, x) K(s, y) ds = \hat{K}''(x, y).$$

The eigenvalues corresponding to these kernels are the same, and they are all positive. Let λ_i^2 ($i = 1, 2, \dots$) be these eigenvalues and $\{\varphi_i(x)\}$ and $\{\psi_i(x)\}$ be the corresponding complete orthonormal systems corresponding to \hat{K}' and \hat{K}'' , respectively. Then we have

$$\lambda_i \int_D K(y, x) \varphi_i(y) dy = \psi_i(x),$$

$$\lambda_i \int_D K(x, y) \psi_i(y) dy = \varphi_i(x).$$

Let $f(x)$ be an arbitrary function such that

$$f(x) = \int_D K(x, y)h(y)dy,$$

where $h(x)$ is a function square integrable on D . The function $f(x)$ can then be expanded in the form

$$f(x) = \sum_{i=1}^{\infty} c_i \varphi_i(x), \quad (8)$$

where

$$c_i = \int_D f(x)\varphi_i(x)dx \quad (i = 1, 2, \dots).$$

The series in the expansion (8) converges uniformly.

The Fredholm integral equation (1) of the first kind with a general kernel (i.e., a kernel that is not necessarily symmetric) has a square integrable solution $\varphi(x)$ if and only if $f(x)$ has a uniformly convergent expansion (8) and $\sum_{i=1}^{\infty} (c_i \lambda_i)^2 < +\infty$. When this condition is satisfied, equation (1) has a solution given by $\sum_{n=1}^{\infty} c_n \lambda_n \varphi_n(x)$ that converges in the sense of †mean convergence.

It should be noted that the theory concerning symmetric kernels can be extended to complex-valued **Hermitian kernels**, i.e., kernels such that $K(x, y) = \overline{K(y, x)}$. Also, we can obtain the theory in Section G and this section, concerning Fredholm integral equations with continuous kernels, by using the methods of functional analysis that treat $\int_D K(x, y)\varphi(y)dy$ as a †compact operator in the space of continuous functions (\rightarrow 68 Compact and Nuclear Operators).

I. Kernels of Hilbert-Schmidt Type

Kernels of Hilbert-Schmidt type are kernels which are square integrable in the sense of Lebesgue over $D \times D$, where D is an arbitrary domain. Most of the results mentioned in the previous section concerning integral equations with kernels continuous on bounded domains are valid also for equations with kernels of Hilbert-Schmidt type, because every operator mentioned in the previous section is also a compact operator in the space concerned, i.e., the space $L_2(D)$ [6] (\rightarrow 68 Compact and Nuclear Operators).

J. Singular Kernels

For general kernels that are not necessarily continuous, the theory described in the previous sections does not apply properly, but when an iterated kernel $K_m(x, y)$ has a resolv-

ent $R_m(x, y)$, we can find a resolvent $R(x, y)$ of $K(x, y)$ in the form

$$R(x, y) = R_m(x, y) + H_m(x, y) + \int_D R_m(x, s)H_m(s, y)ds,$$

where $H_m(x, y) = \sum_{i=1}^{m-1} K^i(x, y)$. When a kernel is of the form $\lambda K(x, y)$, the relationship between eigenvalues and eigenfunctions corresponding to an iterated kernel and those corresponding to the original kernel, which was stated in Sections G and H, is still valid for general kernels. If a kernel $K(x, y)$ is continuous for $x \neq y$ and has a singularity of the form $|x - y|^{-\alpha}$ ($0 < \alpha < 1$) on $x = y$, the iterated kernels $K_m(x, y)$ are continuous provided that $(1 - \alpha)m \geq 1$. †Green's functions of partial differential equations of elliptic type have this property.

A kernel that is not square integrable is called a **singular kernel**. An integral equation whose domain of definition is unbounded or whose kernel is singular is called a **singular integral equation** [9]. Singular integral equations have some particular properties that are not seen in ordinary integral equations, i.e., integral equations with kernels continuous in a bounded closed domain. For example [10], consider the identity

$$\begin{aligned} & \sqrt{\frac{2}{\pi}} \int_0^{\infty} \sin xy \left(\sqrt{\frac{\pi}{2}} e^{-xy} + \frac{y}{\alpha^2 + y^2} \right) dy \\ &= \sqrt{\frac{\pi}{2}} e^{-\alpha x} + \frac{x}{\alpha^2 + x^2}, \end{aligned}$$

where α is an arbitrary real number. This equality shows that for the continuous kernel $\sqrt{2/\pi} \sin xy$, 1 is an eigenvalue and $\sqrt{\pi/2} e^{-\alpha x} + x/(\alpha^2 + x^2)$ is a corresponding eigenfunction. Since α is arbitrary, the index of the eigenvalue 1 is evidently infinity. As another example, observe the equality

$$\int_{-\infty}^{\infty} e^{-|x-y|} e^{-ixy} dy = \frac{2}{1 + \alpha^2} e^{-ixx},$$

where α is an arbitrary real number. From this equality, we see that for the continuous kernel $e^{-|x-y|}$ defined on $(-\infty, \infty)$, and number $\lambda = (1 + \alpha^2)/2$ greater than or equal to 1/2 is an eigenvalue. In this example, the **spectrum**, i.e., the set of eigenvalues, is a continuum. Such a spectrum is called a **continuous spectrum**.

In applications, an important role is played by integral equations with **kernels of Carleman type**:

$$K(x, y) = G(x, y)/(y - x),$$

where $G(x, y)$ is a bounded function. In integral equations with such kernels, the integral

is taken in the sense of the \dagger Cauchy principal value [5, 9, 11, 12]. For example, the **Riemann-Hilbert problem** is the following: We are given a simple closed and smooth curve L in the complex plane and real-valued smooth functions a , b , and c defined on L , $a + ib$ never vanishing. The problem is to find a function $\varphi(z)$ which is holomorphic in the exterior of L , at most of polynomial growth at infinity, and continuous up to L with boundary value φ^+ such that $\operatorname{Re}[(a + ib)\varphi^+] = c$ on L . This problem is reduced to that of finding a function μ defined on L satisfying an integral equation

$$\mu(z) - \int_L \frac{G(z, \zeta)}{z - \zeta} \mu(\zeta) d\zeta = f(z) \quad (z \in L),$$

where G is a smooth kernel determined by $a + ib$, f is a known function depending on c , and the integral is taken in the sense of the Cauchy principal value. The integer κ defined by $\kappa = (1/\pi) \int_L d(\arg(a - ib))$ is called the **index** of this problem. The full solution of the Riemann-Hilbert problem was given by I. N. Vekua [11].

A multidimensional analog of the Cauchy principal value is the **singular integral** of A. P. Calderón and A. Zygmund (\rightarrow 251 Linear Operators). A smooth function $k(x)$ defined in \mathbf{R}^n except at $x = 0$ is called a **kernel of Calderón-Zygmund type** if $k(x)$ is positively homogeneous of order $-n$ and if its integral mean on the unit sphere is zero. Then the operator K defined by

$$Kf(x) = \lim_{\varepsilon \rightarrow 0} \int_{|y| > \varepsilon} k(y) f(x - y) dy$$

is called a Calderón-Zygmund **singular integral operator** [13]. K is a bounded linear operator in $L^p(\mathbf{R}^n)$ if $1 < p < +\infty$. If $n = 1$ and $k(x) = 1/(\pi ix)$, K is nothing but the \dagger Hilbert transformation. The pseudodifferential operator (\rightarrow 345 Pseudodifferential Operators) is an extension in some sense of the singular integral operator.

K. Systems of Integral Equations

A system of Fredholm integral equations of the second kind can always be reduced to a single equation. In fact, as is seen easily, a system of integral equations

$$\varphi_i(x) - \lambda \sum_j \int_0^1 K_{ij}(x, y) \varphi_j(y) dy = f_i(x) \quad (i = 1, 2, \dots, n)$$

can be reduced to a single equation

$$\Phi(x) - \lambda \int_0^n K(x, y) \Phi(y) dy = F(x) \quad (0 \leq x \leq n),$$

where

$$\Phi(x) = \varphi_i(x - i + 1), \quad F(x) = f_i(x - i + 1),$$

$$K(x, y) = K_{ij}(x - i + 1, y - j + 1)$$

$$(i - 1 \leq x, j - 1 \leq y \leq j; i, j = 1, 2, \dots, n).$$

A system of Volterra integral equations of the second kind can be reduced to a single equation by eliminating the unknown functions successively.

L. Integral Equations of Volterra Type

Consider a Volterra integral equation of the first kind

$$\int_a^x K(x, y) \varphi(y) dy = f(x)$$

such that $K(x, x) \neq 0$ and $K_x(x, y)$ and $f'(x)$ are continuous. If we differentiate both sides of the equation, then we have a Volterra integral equation of the second kind:

$$\varphi(x) + \int_a^x \frac{K_x(x, y)}{K(x, x)} \varphi(y) dy = \frac{f'(x)}{K(x, x)}.$$

Abel's integral equation of general form is

$$\int_0^x \frac{G(x, y)}{(x - y)^\alpha} \varphi(y) dy = f(x) \quad (0 < \alpha < 1). \quad (9)$$

If G , G_x , and f' are continuous and $G(x, x) \neq 0$, equation (9) can be reduced to the equation

$$\int_0^u H(u, y) \varphi(y) dy = \int_0^u f(x) (u - x)^{\alpha-1} dx,$$

where

$$H(u, y) = \int_y^u \frac{G(x, y) dx}{(u - x)^{1-\alpha} (x - y)^\alpha}.$$

Since $H(u, u) = (\pi/\sin \alpha\pi) G(u, u) \neq 0$, it follows that

$$\varphi(u) + \int_0^u \frac{H_u(u, y)}{H(u, u)} \varphi(y) dy = g(u), \quad (9a)$$

where

$$\begin{aligned} g(u) &= H(u, u)^{-1} \frac{d}{du} \int_0^u f(x) (u - x)^{\alpha-1} dx \\ &= H(u, u)^{-1} \left(u^{\alpha-1} f(0) \right. \\ &\quad \left. + \int_0^u (u - x)^{\alpha-1} f'(x) dx \right). \end{aligned}$$

Clearly (9a) is a Volterra integral equation of the second kind. **Abel's problem** (\rightarrow Section A) was to find the path of a falling body for a given time of descent. The problem then can be reduced to the solution of equation (9) with $G(x, y) \equiv 1$ and $\alpha = 1/2$. When $G(x, y) \equiv 1$, we

can solve equation (9) explicitly to get

$$\varphi(x) = \frac{\sin \alpha \pi}{\pi} \left(x^{\alpha-1} f(0) + \int_0^x (x-t)^{\alpha-1} f'(t) dt \right)$$

[5].

M. Nonlinear Integral Equations

When a **nonlinear integral equation** includes a parameter λ , it may happen that the parameter has a **bifurcation point**, i.e., a value λ_0 such that the number of real solutions is changed when λ varies through λ_0 taking real values. For example, consider the equation

$$\varphi(x) - \lambda \int_0^1 \varphi^2(y) dy = 1.$$

This equation has real solutions $\varphi(x) = (1 \pm \sqrt{1-4\lambda})/2\lambda$ for $\lambda \leq 1/4$ but no real solution for $\lambda > 1/4$. Hence $\lambda_0 = 1/4$ is a bifurcation point (\rightarrow 286 Nonlinear Functional Analysis).

Among nonlinear integral equations, **Hammerstein's integral equation** has been studied in detail [5, 14]. It is an equation of the form

$$\varphi(x) + \int_D K(x, y) f(y, \varphi(y)) dy = 0. \quad (10)$$

If $K(x, y)$ and $f(y, 0)$ are square integrable and $f(y, u)$ satisfies a Lipschitz condition in u with a sufficiently small coefficient, then the integral equation (10) can be solved by successive approximations. If $K(x, y)$ is a square integrable positive kernel and $\int_D |K(x, y)|^2 dy$ is bounded, then we can prove the existence and uniqueness of a solution of (10) under a condition on $f(y, u)$ weaker than a Lipschitz condition. We can prove similar results for equation (10) with a nonsymmetric kernel when $K(x, y)$ is **continuous in the mean**, that is,

$$\lim_{x' \rightarrow x} \int_D |K(x', y) - K(x, y)|^2 dy = 0,$$

$$\lim_{y' \rightarrow y} \int_D |K(x, y') - K(x, y)|^2 dx = 0.$$

A nonlinear Volterra integral equation of the form

$$\varphi(x) = f(x) + \int_a^x F(x, y, \varphi(y)) dy$$

can be solved by successive approximations if $F(x, y, u)$ and $f(x)$ are square integrable, $F(x, y, u)$ satisfies a Lipschitz condition $|F(x, y, u') - F(x, y, u'')| \leq k(x, y)|u' - u''|$ with some square integrable function $k(x, y)$, and $\int_a^x F(x, y, f(y)) dy$ is majorized by some square integrable function of x . When $F(x, y, u)$ and $f(x)$ are continuous, we can obtain theorems on the existence and uniqueness of continuous solutions and comparison theorems similar to

those for initial value problems in ordinary differential equations [16].

N. Numerical Solution

For the **numerical solution of integral equations**, we assume throughout that the functions appearing are all continuous and the solution of every equation is unique. Methods of numerical solution can be divided roughly into two classes. Methods of one class try to evaluate numerically the analytical solution described in the preceding sections, and methods of the other try to obtain a solution by transforming the problem to one that is numerically solvable.

(1) A Method Based on Numerical Quadrature.

Consider the integral equation

$$\int_a^b F(x, y, \varphi(x), \varphi(y)) dy = 0.$$

Let $a = x_1 < x_2 < \dots < x_n = b$ be points on the interval $[a, b]$ and $\varphi_1, \varphi_2, \dots, \varphi_n$ be the values of $\varphi(x)$ at x_1, x_2, \dots, x_n . By the use of numerical quadrature, we then have the following system of equations in φ_i :

$$\sum_{j=1}^n a_j F(x_i, x_j, \varphi_i, \varphi_j) = 0 \quad (i = 1, 2, \dots, n).$$

The method corresponds to that of solving ordinary differential equations by their difference equation analogs. Hence the errors involved in the solutions obtained by this method can be analyzed similarly to those in the case of numerical solution of ordinary differential equations (\rightarrow 303 Numerical Solution of Ordinary Differential Equations). If the given integral equation is a Fredholm equation of the second kind, then we have a system of linear equations in φ_i . If we apply quadrature formulas to the integral appearing in the integral equation by using the values of φ_i obtained, then we have a formula by which the solution can be evaluated directly, that is, without using an interpolation formula.

(2) A Method Utilizing Recurrence Formulas.

Let d_n and $d_n(x, y)$ be the respective coefficients of λ^n in the expansions of Fredholm's determinant $D(\lambda)$ and Fredholm's first minor $D(x, y; \lambda)$. They satisfy the recurrence formulas

$$d_n(x, y) = d_n K(x, y) + \int_a^b K(x, s) d_{n-1}(s, y) ds,$$

$$d_{n+1} = -\frac{1}{n+1} \int_a^b d_n(s, s) ds,$$

$$d_0 = 1, \quad d_0(x, y) = K(x, y).$$

By the use of these formulas, we can compute

d_n and $d_n(x, y)$ successively and hence evaluate $D(\lambda)$ and $D(x, y; \lambda)$ approximately, and by means of these recurrence formulas we can readily obtain a solution of a Fredholm equation of the second kind.

(3) A Method Utilizing Approximate Kernels.

If we replace a kernel by an approximate one in a Fredholm integral equation of the second kind, then we have an integral equation that has a solution approximately equal to the solution of the original equation. Hence if we can find an approximate kernel for which an integral equation can be solved numerically or analytically, then we can find an approximation to the desired solution by solving the modified equation. For such solutions, a method of error estimation was given by F. G. Tricomi [17].

(4) An Iterative Method. Consider the integral equation

$$\varphi(x) = \int_a^b F(x, y, \varphi(x), \varphi(y)) dy.$$

Let $\varphi_0(x)$ be an adequate function, and define $\varphi_n(x)$ successively by

$$\varphi_{n+1}(x) = \int_a^b F(x, y, \varphi_n(x), \varphi_n(y)) dy.$$

If the sequence $\{\varphi_n(x)\}$ converges, then the limit $\lim_{n \rightarrow \infty} \varphi_n(x) = \varphi(x)$ is a solution of the given equation, and hence we can obtain an approximation to a solution by calculating $\varphi_n(x)$ for some finite n . This method can be used effectively for Fredholm integral equations of the second kind with a parameter λ , provided that the absolute value of λ is smaller than the least absolute value of the eigenvalues.

(5) Variational Method. If some conditions are fulfilled, an integral equation of the form

$$G(x, \varphi(x)) + \int_a^b F(x, y, \varphi(x), \varphi(y)) dy = 0$$

can be regarded as an Euler-Lagrange equation for a variational problem

$$J[u] = \int_a^b \int_a^b E(x, y, u(x), u(y)) dx dy + \int_a^b H(x, u(x)) dx = \text{extremal.} \quad (11)$$

In this case, we can find a solution of the given integral equation numerically by solving the variational problem (11) numerically [18].

(6) Enskog's Method. Suppose that $\{\varphi_n(x)\}$ is a complete orthonormal system for the Fred-

holm integral equation (7). If we put $\psi_n(x) = \varphi_n(x) - \lambda \int_a^b K(y, x) \varphi_n(y) dy$, then from (7) we have

$$\int_a^b \varphi(x) \psi_n(x) dx = \int_a^b f(x) \varphi_n(x) dx, \quad (12)$$

and furthermore we see that $\{\psi_n(x)\}$ can be orthonormalized to yield a complete orthonormal system $\{\chi_n(x)\}$. The equality (12) then shows that the Fourier coefficients of a solution $\varphi(x)$ with respect to the system $\{\chi_n(x)\}$ can be obtained readily from the Fourier coefficients of $f(x)$ with respect to the system $\{\varphi_n(x)\}$. This method of obtaining a solution is called **Enskog's method**.

For Volterra integral equations, methods (1) and (4) can be used effectively. We usually transform equations of the first kind into equations of the second kind by differentiation and then apply the above numerical methods. This is done for the sake of securing stability of the numerical methods.

References

- [1] I. Fredholm, Sur une classe d'équations fonctionnelles, *Acta Math.*, 27 (1903), 365–390.
- [2] G. Vivanti and F. Schwank, *Elemente der Theorie der Lineare Integralgleichungen*, Helwingsche Verlagsbuchhandlung, 1929.
- [3] E. Hellinger and O. Toeplitz, *Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten*, Teubner, 1928.
- [4] R. Courant and D. Hilbert, *Methods of mathematical physics I, II*, Interscience, 1953, 1962.
- [5] F. G. Tricomi, *Integral equations*, Interscience, 1957.
- [6] K. Yosida, *Lectures on differential and integral equations*, Interscience, 1960. (Original in Japanese, 1950.)
- [7] E. Goursat, *Cours d'analyse mathématique III*, Gauthier-Villars, 1928.
- [8] F. Smithies, *Integral equations*, Cambridge, 1958.
- [9] T. Carleman, *Sur les équations intégrales singulières à noyau réel et symétrique*, Uppsala, 1923.
- [10] T. Lalesco, *Introduction à la théorie des équations intégrales*, Gauthier-Villars, 1912.
- [11] N. I. Muskhelishvili, *Singular integral equations*, Noordhoff, 1953. (Original in Russian, 1946.)
- [12] S. G. Mikhlin, *Multidimensional singular integrals and integral equations*, Pergamon, 1965. (Original in Russian, 1962.)
- [13] A. P. Calderón and A. Zygmund, *Singular integral operators and differential equations*, *Amer. J. Math.*, 79 (1957), 901–921.

[14] H. Schaefer, *Neue Existenzsätze in der Theorie nichtlinearer Integralgleichungen*, Akademie-Verlag, 1955.
[15] L. Lichtenstein, *Vorlesungen über einige Klassen nichtlinearer Integralgleichungen und Integro-differentialgleichungen*, Springer, 1931.
[16] T. Sato, *Sur l'équation intégrale non linéaire de Volterra*, *Compositio Math.*, 11 (1953), 271–290.
[17] H. Büchner, *Die praktische Behandlung von Integralgleichungen*, Springer, 1952.
[18] L. Collatz, *The numerical treatment of differential equations*, Springer, 1960.
[19] L. V. Kantorovich and V. I. Krylov, *Approximate methods of higher analysis*, Interscience, 1958. (Original in Russian, 1949.)

218 (VII.19)
Integral Geometry

A. General Remarks

Integral geometry, in the broad sense, is the branch of geometry concerned with integrals on manifolds, but the problems considered in **integral geometry** are usually of a more limited nature. If a \dagger Lie group G acts on a \dagger differentiable manifold M as a \dagger Lie transformation group, G also acts on various figures on M , by which we mean geometric objects such as \dagger submanifolds of M , \dagger tangent r -frame bundles on M , etc. Let \mathcal{F} be a set of such figures on M invariant under G (i.e., $gF \in \mathcal{F}$ for $g \in G, F \in \mathcal{F}$). Consider the following problems: (i) to know whether any G -invariant \dagger measure μ on \mathcal{F} exists, and how to determine μ if it exists; (ii) to find the integral $\int \varphi(F) d\mu(F)$ of functions φ on \mathcal{F} with respect to the measure μ .

The term **integral geometry** was introduced by W. Blaschke, who considered the special case of problem (ii) in which $\varphi(F)$ is a function representing geometric properties of F and the integral is to be evaluated by means of the geometric invariants concerning \mathcal{F} [1]. Problems of so-called **geometric probability** (such as the problem of Buffon's needle) belong to this category. The measure μ is called the **kinetic measure** (or **kinetic density**), and $d\mu(F)$ is also denoted by dF . If \mathcal{F} has the structure of an n -dimensional differentiable manifold and the measure μ is given by a \dagger volume element ω (i.e., a positive \dagger differential form of degree n), we denote ω also by dF . Problem (i) is simple: If G acts \dagger transitively on \mathcal{F} , then $\mathcal{F} = G/H$, where H is the \dagger isotropy subgroup of G . In this case \mathcal{F} has the structure of a differentiable manifold, and if a G -invariant \dagger (Radon) measure exists, it is unique up to a multiplicative

constant. A condition for the existence of a G -invariant measure μ can be given by means of \dagger Haar measures of G and H (\rightarrow 225 Invariant Measures). We now consider some examples.

B. Crofton's Formula

Let $G(p, \theta)$ be a straight line defined by the equation $x_1 \cos \theta + x_2 \sin \theta = p$ with respect to orthogonal coordinates in a Euclidean plane. Let $n(p, \theta)$ be the number of intersections of $G(p, \theta)$ with a curve C of length L . Then we have **Crofton's formula**,

$$\int n(p, \theta) dp d\theta = 2L, \tag{1}$$

where $dp d\theta$ is the \dagger exterior product of the differential forms $dp, d\theta$ of degree 1, and the integral is extended over $p \in (-\infty, \infty)$ and $\theta \in [0, 2\pi]$.

C. Poincaré's Formula and the Principal Formula of Integral Geometry

The kinetic density dF of a figure F congruent (with the same orientation) to a fixed figure in a Euclidean plane is defined as follows: Let R be an orthogonal frame attached to F , (x_1, x_2) be the coordinates of the origin of R with respect to a fixed orthogonal frame R_0 , and θ be the angle between the first axis of R and the first axis of R_0 . If we put $dF = dx_1 dx_2 d\theta$ (exterior product), dF has the following invariance properties: (i) dF is not changed by displacements of F ; (ii) dF is not changed if instead of R we take another orthogonal frame R' attached to F .

Let two plane curves C_1, C_2 of length L_1, L_2 , respectively, be given, and suppose that C_1 is fixed and C_2 is mobile. If the number of intersections of C_2 in an arbitrary position with C_1 is finite and equal to n , then the integral of n extended over all possible positions of C_2 is given by

$$\int n dC_2 = 4L_1 L_2 \tag{2}$$

(**Poincaré's formula**). L. A. Santaló applied this result to give a solution of the \dagger isoperimetric problem (1936).

Let C_1, C_2 be two plane \dagger Jordan curves of length L_1, L_2 , respectively, and let S_1, S_2 be the areas of the domains bounded by C_1, C_2 , respectively. Suppose that C_1 is fixed and C_2 mobile, and let χ be the number of connected domains common to the domains bounded by C_1, C_2 for C_2 in an arbitrary position. Then the integral of χ extended over all possible

positions of C_2 intersecting C_1 is given by

$$\int \chi dC_2 = L_1 L_2 + 2\pi(S_1 + S_2) \quad (3)$$

(Blaschke [1]). This is the **principal formula of integral geometry**. Many formulas can be derived from it as special cases or limiting cases.

D. Generalization to Dimension n

The kinetic density of subspaces of dimension k in a Euclidean space and in a spherical space of dimension n was given by Blaschke, while the generalization of the principal formula (3) to a Euclidean space of dimension n was given by S. S. Chern, applying the methods of E. Cartan.

Let (e_1, \dots, e_n) be a positively oriented orthonormal frame with vertex A . The infinitesimal relative displacements are then given by

$$dA = \sum_{i=1}^n \omega_i e_i, \quad de_i = \sum_{j=1}^n \omega_{ij} e_j,$$

where $\omega_i = (dA, e_i)$, $\omega_{ij} = (de_i, e_j) = -\omega_{ji}$ are differential forms of degree 1 in the orthogonal coordinates of A and the $n(n-1)/2$ variables that determine e_1, \dots, e_n . For various positions of a figure, we take an orthogonal frame (A, e_1, \dots, e_n) fixed to this figure and form the exterior product

$$dK = \bigwedge_i \omega_i \bigwedge_{i < j} \omega_{ij}$$

of all the ω_i and ω_{ij} ($i < j$). This has the invariance properties (i) and (ii) of Section C and is, by definition, the kinetic density of the figure in an n -dimensional Euclidean space.

Moreover, the kinetic density of dE of k -dimensional subspaces E can be obtained by considering the orthogonal frames such that the vertex A and e_1, \dots, e_k lie on E and by forming the exterior product of the corresponding ω_α , $\omega_{\alpha\lambda}$ ($\alpha = 1, \dots, k; \lambda = k+1, \dots, n$),

$$dE = \bigwedge \omega_\alpha \bigwedge \omega_{\alpha\lambda}.$$

Let Σ be a compact orientable hypersurface of class C^2 in an n -dimensional Euclidean space, and let k_α ($\alpha = 1, \dots, n-1$) be the principal curvatures at a point on Σ . Denote by S_i the elementary symmetric form of degree i in k_α ($i = 1, \dots, n-1$), and put $S_0 = 1$. Then consider the integrals over Σ :

$$M_i = \int_\Sigma S_i dS / \binom{n-1}{i}, \quad i = 0, 1, \dots, n-1, \quad (4)$$

where dS denotes the surface element of Σ . Let D_1, D_2 be the domains bounded by two compact orientable hypersurfaces Σ_1, Σ_2 of class C^2 with volume V_1, V_2 , respectively, and let

$M_i^{(1)}, M_i^{(2)}$ be the integrals M_i defined by (4) for Σ_1, Σ_2 . If Σ_1 is fixed, Σ_2 is mobile, and the Euler-Poincaré characteristic χ of $D_1 \cap D_2$ is finite, then the generalization of (3) has the form

$$\int \chi d\Sigma_2 = I_1 \dots I_{n-1} \left(M_{n-1}^{(1)} V_2 + M_{n-1}^{(2)} V_1 + \frac{1}{n} \sum_{k=0}^{n-2} M_k^{(1)} M_{n-2-k}^{(2)} \right) \quad (5)$$

(Chern's formula), where $d\Sigma_2$ is the kinetic density of Σ_2 and I_k ($k = 1, \dots, n-1$) is the area of the unit sphere in a Euclidean space of dimension $k+1$, with the integral extended over all positions of Σ_2 intersecting Σ_1 .

Let dE be the kinetic density of the subspaces E of dimension k intersecting a compact orientable hypersurface Σ of class C^2 , and let χ be the Euler characteristic of the intersection of E with the domain bounded by Σ . The integral $\int \chi dE$ extended over all hyperplanes of dimension k intersecting Σ is proportional to M_k relative to the hypersurface Σ defined by (4). This fact generalizes (1). Further generalizations were obtained by Chern (1966).

E. Other Generalizations

For 2-dimensional spaces of constant curvature, Santaló derived formulas analogous to those in a Euclidean plane (1942–1943) and thus solved the isoperimetric problem in these spaces. In 1952, he derived a formula corresponding to (5) in n -dimensional spaces of constant curvature, following Chern's method. He investigated further integral geometry in affine, projective, and Hermitian spaces.

Chern and others obtained the results of the previous sections by Cartan's method of general moving frames and studied integral geometry in the setting of the geometry of Lie transformation groups in the sense of F. Klein (\rightarrow 137 Erlangen Program).

Chern, P. Griffiths, and others studied the value distribution theory of holomorphic mappings in several complex variables from the point of view of integral geometry (1961) (\rightarrow 21 Analytic Functions of Several Complex Variables, 124 Distribution of Values of Functions of a Complex Variable).

F. Radon Transforms

Another important topic of integral geometry is the theory of Radon transforms. Let \mathcal{F} be the set of hyperplanes $\xi(\omega, p) = \{x \in \mathbf{R}^n | (x, \omega) = p\}$ in the Euclidean space \mathbf{R}^n , where $\omega = (\lambda_1, \dots, \lambda_n)$ is a unit vector, $(x, \omega) = \sum x_i \lambda_i$, and p is real. For a function f defined in \mathbf{R}^n , define

$\hat{f}(\xi) = \hat{f}(\omega, p)$ on \mathcal{F} by

$$\hat{f}(\xi) = \int_{\xi} f(x) d_{\xi} x, \quad (6)$$

where $d_{\xi} x$ is the \dagger volume element on the hyperplane ξ such that $d_{\xi} x \wedge \sum \lambda_i dx_i = dx$, with dx the volume element of \mathbf{R}^n . Then \hat{f} is called the **Radon transform** of f . For example, if f is the \dagger characteristic function of a bounded domain V , the value $\hat{f}(\xi)$ of the Radon transform \hat{f} of f at $\xi \in \mathcal{F}$ is the volume of the section of V by ξ . Now the group G of \dagger motions of \mathbf{R}^n (the \dagger connected component of the identity of the group of isometries) acts on \mathcal{F} transitively. For every $x \in \mathbf{R}^n$, the \dagger isotropy subgroup G_x of G with respect to x acts transitively on the set $\check{x} = \{\xi \in \mathcal{F} \mid x \in \xi\}$. Since G_x is compact, there exists a unique normalized G_x -invariant measure μ on \check{x} such that $\mu(\check{x}) = 1$. For a function g on \mathcal{F} , the **conjugate Radon transform** \check{g} can now be defined by

$$\check{g}(x) = \int_{x \in \xi} g(\xi) d\mu(\xi) \quad (7)$$

as a function on \mathbf{R}^n .

The determination of \check{g} belongs to problem (ii) of integral geometry mentioned in Section A. In particular, it is important to determine $\check{g} = (\hat{f})^\sim$ for $g = \hat{f}$ and to find the relation between $(\hat{f})^\sim$ and f . These problems were solved by J. Radon for $n = 2, 3$ and by F. John in the general case. The results can be formulated as follows:

In the case of odd n , let \mathcal{S} be the space of \dagger rapidly decreasing C^∞ -functions (\rightarrow 168 Function Spaces). Let $\mathcal{S}^*(\mathcal{F})$ be the set of $g(\omega, p) \in \mathcal{S}(\mathcal{F})$ such that $\int_{-\infty}^{\infty} g(\omega, p) p^k dp = 0$ for every natural number k and every ω . For every $f \in \mathcal{S}(\mathbf{R}^n)$ and every $g \in \mathcal{S}^*(\mathcal{F})$, we then have

$$f = c\Delta^{(n-1)/2}((\hat{f})^\sim), \quad g = cL^{(n-1)/2}((\check{g})^\sim),$$

where Δ is the \dagger Laplacian in \mathbf{R}^n and $Lg(\omega, p) = d^2 g(\omega, p)/dp^2$, $c = \Gamma(n/2)^{-1} (2\pi i)^{1-n} \pi^{n/2}$.

In the case of even n , for every $f \in \mathcal{S}(\mathbf{R}^n)$, $g \in \mathcal{S}^*(\mathcal{F})$,

$$f = c_1 J_1((\hat{f})^\sim), \quad g = c_2 J_2((\check{g})^\sim),$$

where

$$J_1(f)(x) = \int_{\mathbf{R}^n} f(y) |x - y|^{1-2n} dy,$$

$$J_2(g)(\omega, p) = \int_{\mathbf{R}} g(\omega, p) |p - q|^{-n} dq.$$

These integrals are in general divergent, and they must be interpreted as regularizations defined by analytic continuation with respect to the powers of $|x - y|$ or $|p - q|$ [9].

John applied the Radon transform on \mathbf{R}^n to the study of partial differential equations in \mathbf{R}^n

A formula corresponding to \dagger Plancherel's theorem and an analog of the \dagger Paley-Wiener theorem for the Fourier transform are valid for the Radon transform (I. M. Gel'fand et al. [6]).

G. Horospheres

The theory of the Radon transform is also important in noncompact \dagger symmetric Riemannian spaces M . The connected component G of the identity in the group of isometries of M is isomorphic to the \dagger adjoint group $\text{Ad } G$ and can be considered as a linear group. Maximal \dagger unipotent subgroups of G are conjugate to each other. If N is such a subgroup, we call the \dagger orbits on M of gNg^{-1} **horospheres** on M for $g \in G$. These correspond to the hyperplanes in \mathbf{R}^n . If M is the complex upper half-plane with the \dagger hyperbolic non-Euclidean metric, the horospheres are precisely the circles tangent to the real axis and the straight lines parallel to the real axis.

The group G acts transitively on the set \mathcal{F} of horospheres on M , and we have $\mathcal{F} = G/M_0 N$. Here $G = KAN$ is an \dagger Iwasawa decomposition of G , and M_0 is the \dagger centralizer of A in K . For a horosphere $\xi \in \mathcal{F}$, let $d_{\xi} x$ be the volume element on ξ with respect to the \dagger Riemannian metric on ξ induced by the Riemannian metric on M , and define the Radon transform \hat{f} of a function f on M by (6) as a function on \mathcal{F} . For every $x \in M$, there exists a unique normalized measure on $\check{x} = \{\xi \in \mathcal{F} \mid x \in \xi\}$ invariant under the (compact) isotropy subgroup G_x of G at x ($\mu(\check{x}) = 1$). The conjugate Radon transform \check{g} of a function g on \mathcal{F} is defined by formula (7) by means of this measure μ . Then there exists an integrodifferential operator \wedge such that if \wedge^* is the adjoint operator, we have the inversion formula $f = (\wedge \wedge^* \hat{f})^\sim$ and **Plancherel's theorem**:

$$\int_M |f(x)|^2 dx = \int_{\mathcal{F}} |\wedge \hat{f}(\xi)|^2 d\xi,$$

where dx , $d\xi$ are G -invariant measures on M , \mathcal{F} , respectively, and f is an arbitrary C^∞ -function with compact support. If the \dagger Cartan subgroups of G are conjugate to each other, \wedge is a differential operator; the inversion formula can then be written in the form $f = L((\hat{f})^\sim)$ with some differential operator L on M [9].

S. Helgason applied the Radon transform on noncompact symmetric Riemannian spaces to solve differential equations on these spaces (1973).

Horospheres and Radon transforms can be defined not only for symmetric Riemannian spaces G/K , but also for various \dagger homoge-

$x_k \leq b_k, k=1, 2, \dots, n\}$ ($-\infty \leq a_k < b_k \leq \infty$) is defined by $m(I) = \prod_{k=1}^n (b_k - a_k)$. Let \mathfrak{M}_0 be the collection of all sets that can be represented as the finite union of disjoint left-open intervals, and for such expression $A = \bigcup_{j=1}^r I_j$ in \mathfrak{M}_0 define $m(A) = \sum_{j=1}^r m(I_j)$. Let $\mathfrak{M} = \mathfrak{M}_0 \cup \{\emptyset\}$ and $m(\emptyset) = 0$. Then m gives a finitely additive measure on \mathfrak{M} that is completely additive. Therefore m determines an outer measure μ^* , which in turn determines a measure μ . This μ^* is called the **Lebesgue outer measure** (or simply **outer measure**). Sets that are measurable with respect to μ^* are said to be **Lebesgue measurable** (or simply **measurable**), and the measure μ is called the (n -dimensional) **Lebesgue measure** (or simply **measure**). Every interval is measurable, and its measure coincides with its volume. Open sets, closed sets, and Borel sets are all measurable. More generally, suppose that an outer measure μ^* defined on a metric space X with a metric d satisfies the condition: If $d(A, B) > 0$, then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$. Then every closed subset of X is μ^* -measurable, and therefore so is every Borel subset. Here $d(A, B) = \inf\{d(a, b) | a \in A, b \in B\}$ denotes the distance between the two sets A and B . A measure μ defined on the class of all Borel subsets of a topological space X is called a **Borel measure**. The cardinality of the set of all Lebesgue measurable subsets of \mathbf{R}^n is 2^c , while the cardinality of the class of Borel sets is c (here c is the cardinal number of the t continuum, i.e., the cardinal number of \mathbf{R}). Therefore there exists a Lebesgue measurable set that is not a Borel set. It follows immediately from the definition that the Lebesgue measure is K -regular, so for every Lebesgue measurable set A we can find an $^tF_\sigma$ -set B and a $^tG_\delta$ -set C such that $B \subset A \subset C$ and $\mu(C - B) = 0$.

Historically, for a bounded subset A of \mathbf{R}^n , C. Jordan defined $\bar{m}(A)$ to be the infimum of all possible values $m(B)$, where $B \in \mathfrak{M}$ and $B \supset A$, and $\underline{m}(A)$ to be supremum of all possible values $m(B)$, where $B \in \mathfrak{M}$ and $B \subset A$. He called $\bar{m}(A)$ the **outer volume** of A and $\underline{m}(A)$ the **inner volume** of A (in the case of \mathbf{R}^2 , the **outer area** and **inner area**, respectively). When $\bar{m}(A) = \underline{m}(A)$, A is called **Jordan measurable**, and this common value is defined to be the **Jordan measure** (**Jordan content**) of A . Jordan measure is only finitely additive, and was found to be unsatisfactory in many respects. It was Lebesgue who modified this notion and introduced completely additive measures. Jordan measurable sets are always Lebesgue measurable.

Using the t axiom of choice, a set that is not Lebesgue measurable can be constructed (G. Vitali). For example, a set obtained by choosing exactly one element from each coset of the additive group of all rationals in the addi-

tive group of the reals is not measurable [3, pp. 67–70].

H. Product Measure

When two σ -finite complete measures μ_X and μ_Y are defined on completely additive classes \mathfrak{B}_X and \mathfrak{B}_Y on X and Y , respectively, an element C belonging to the smallest finitely additive class \mathfrak{R} in the Cartesian product that contains $\{A \times B | A \in \mathfrak{B}_X, B \in \mathfrak{B}_Y\}$ can be represented as a finite disjoint union $C = \bigcup_{j=1}^n (A_j \times B_j)$ ($A_j \in \mathfrak{B}_X, B_j \in \mathfrak{B}_Y$). If we define $v(C) = \sum_{j=1}^n \mu_X(A_j)\mu_Y(B_j)$ (here we agree to put $0 \cdot \infty = 0$), then this value is independent of the way the set C is represented, and v defines a completely additive measure on \mathfrak{R} . By extending v by means of Hopf's extension theorem, we obtain a (complete) measure space, called the **(complete) product measure space** obtained from the measure spaces $(X, \mathfrak{B}_X, \mu_X)$ and $(Y, \mathfrak{B}_Y, \mu_Y)$. The measure obtained in this way on the space $X \times Y$ is called the **product measure** of μ_X and μ_Y and is denoted by $\mu_X \times \mu_Y$. If we denote by \mathfrak{M}_p the class of all Lebesgue measurable subsets of the p -dimensional Euclidean space \mathbf{R}^p and by m_p the p -dimensional Lebesgue measure, then the (complete) product measure space of $(\mathbf{R}^p, \mathfrak{M}_p, m_p)$ and $(\mathbf{R}^q, \mathfrak{M}_q, m_q)$ is $(\mathbf{R}^{(p+q)}, \mathfrak{M}_{p+q}, m_{p+q})$. The product measure space of any finite number of measure spaces $(X_i, \mathfrak{B}_i, \mu_i)$ ($i = 1, \dots, n$) is defined similarly.

Let X_λ ($\lambda \in \Lambda$) be spaces with an index set of arbitrary cardinality. For the product space $X = \prod_{\lambda \in \Lambda} X_\lambda$, an n -**cylinder set**, or simply a **cylinder set**, is a set of the form $A \times \prod_{\lambda \notin \{\lambda_1, \dots, \lambda_n\}} X_\lambda$ ($A \subset X_{\lambda_1} \times \dots \times X_{\lambda_n}$). If a finitely additive class \mathfrak{A}_λ is given for each X_λ , the class of all subsets that can be represented as the union of a finite number of cylinder sets of the form $A_1 \times A_2 \times \dots \times A_n \times \prod_{\lambda \notin \{\lambda_1, \dots, \lambda_n\}} X_\lambda$ ($A_j \in \mathfrak{A}_{\lambda_j}, j = 1, 2, \dots, n$) is a finitely additive class in the space X .

When each \mathfrak{A}_λ is completely additive, the completely additive class \mathfrak{A} generated by this finitely additive class is called the **product** of the completely additive classes \mathfrak{A}_λ and is denoted by $\prod_{\lambda \in \Lambda} \mathfrak{A}_\lambda$. When a measure space $(X_\lambda, \mathfrak{B}_\lambda, \mu_\lambda)$ with $\mu_\lambda(X_\lambda) = 1$ is given for $\lambda \in \Lambda$, a measure μ can be defined in the following way on the completely additive class $\mathfrak{B} = \prod_{\lambda \in \Lambda} \mathfrak{B}_\lambda$ in the product space X : To begin with, for a cylinder set of the form $A_1 \times \dots \times A_n \times X'$ (here $A_j \in \mathfrak{B}_{\lambda_j}$ and $X' = \prod_{\lambda \notin \{\lambda_1, \dots, \lambda_n\}} X_\lambda$), we define $\mu(A_1 \times \dots \times A_n \times X') = \mu_{\lambda_1}(A_1) \dots \mu_{\lambda_n}(A_n)$. If we extend μ to the finitely additive class \mathfrak{C} consisting of all sets that can be represented as the finite union of such cylinder sets, then this extension gives a completely additive measure on \mathfrak{C} , and therefore, by Hopf's extension theorem, there exists a unique extension to a mea-

sure μ on \mathfrak{B} , and μ satisfies $\mu(X) = 1$. We denote this μ by $\mu = \prod_{\lambda \in \Lambda} \mu_\lambda$.

I. Radon Measure

Let X be a * locally compact Hausdorff space, \mathfrak{B} the topological σ -algebra on X , and $C_0(X)$ the real linear space of all real-valued continuous functions f on X having * compact support (i.e., the closure of the set $\{x | f(x) \neq 0\}$ is compact). A (real) * linear functional ϕ defined on $C_0(X)$ is called a **positive Radon measure** if $\phi(f) \geq 0$ whenever $f \geq 0$. For such a functional ϕ there corresponds a Borel measure μ on X for which $\phi(f) = \int_X f d\mu$ holds for every $f \in C_0(X)$. Lebesgue measure m_n is regarded as a positive Radon measure on \mathbf{R}^n . If X is σ -compact (i.e., X can be represented as the countable union of compact sets), then the property $\phi(f) = \int_X f d\mu$ for all $f \in C_0(X)$ defines the measure μ uniquely on the class of Borel sets of X . A linear functional on $C_0(X)$ that can be written as the difference of two positive Radon measures is called a **Radon measure**. For a linear functional ϕ defined on $C_0(X)$ to be a Radon measure, it is necessary and sufficient that for an arbitrary $f \in C_0(X)$, the set $\{\phi(g) | |g| \leq |f|, g \in C_0(X)\}$ be bounded. Equivalently, it is necessary and sufficient that the restriction of ϕ to the subspace of all functions in $C_0(X)$ having their support in a fixed compact subset of X must be continuous with respect to the * topology of uniform convergence. Therefore, if X is compact, an arbitrary continuous linear functional on $C(X)$ is a Radon measure. L. Schwartz investigated Radon measures on spaces that are not locally compact [6].

J. Measurable Functions

When a completely additive class \mathfrak{B} on a space X is given, a function f defined on a \mathfrak{B} -measurable set E and taking real (and possibly $\pm\infty$) values is called a **\mathfrak{B} -measurable function** on E if for an arbitrary real number α , the set $\{x | f(x) > \alpha\}$ is \mathfrak{B} -measurable. The condition $f > \alpha$ may be replaced by $f \geq \alpha$, $f < \alpha$, $f \leq \alpha$. A function f can also be defined to be \mathfrak{B} -measurable if the * inverse image under f of any Borel set is \mathfrak{B} -measurable. When f and g are \mathfrak{B} -measurable, so are $af + bg$ (a, b constants), $f \cdot g$, f/g , $\max(f, g)$, $\min(f, g)$, $|f|^p$ (p a constant), whenever they are well defined. The superior and inferior limits of a sequence of \mathfrak{B} -measurable functions are also \mathfrak{B} -measurable. In a complete product measure space of two σ -finite measure spaces, a function $f(x, y)$ may fail to be measurable as a function of two variables even if it is measurable with respect

to each of the variables x and y separately. For example, let \mathfrak{F} be the class of all closed subsets F of \mathbf{R}^2 satisfying $m_2(F) > 0$, $F \subset [0, 1] \times [0, 1]$. Using the fact that the cardinality of \mathfrak{F} does not exceed the cardinal number of the continuum, we can prove by * transfinite induction that the elements of \mathfrak{F} can be indexed as F_ξ , $\bar{\xi} < \bar{c}$, $\bar{\xi}$ being the cardinality of ξ , and that for every $F_\xi \in \mathfrak{F}$ we can pick two points $z_\xi = (x_\xi, y_\xi)$, $z'_\xi = (x'_\xi, y'_\xi)$ in such a way that if $F_\xi \neq F_\eta$, then $x_\xi, x'_\xi, x_\eta, x'_\eta$ are all distinct and $y_\xi, y'_\xi, y_\eta, y'_\eta$ are all distinct. Furthermore, we can prove that the set E consisting of all such z_ξ is not measurable. Therefore, denoting the characteristic function of the set E by $f(x, y)$, $f(x, y)$ is not measurable; but if we fix x (resp. y), then as a function of y (resp. x), $f(x, y)$ is measurable, since $f(x, \cdot)$ (resp. $f(\cdot, y)$) is always 0 except possibly at one point. If \mathfrak{B} is the class of all Lebesgue measurable sets or the class of all Borel sets, then a \mathfrak{B} -measurable function is called a **Lebesgue measurable function** (or simply **measurable function**) or a **Borel measurable function**, respectively.

In Euclidean space, the class of all Borel measurable functions coincides with the class of all Baire functions, and an arbitrary Lebesgue measurable function is equal almost everywhere to a Baire function of at most the second class. For a function f that is finite almost everywhere on a Lebesgue measurable set E to be Lebesgue measurable, it is necessary and sufficient that for an arbitrary $\varepsilon > 0$ we can find a closed subset F such that $m(E - F) < \varepsilon$ and f is continuous on F (**Luzin's theorem**).

If a sequence $\{f_n\}$ of \mathfrak{B} -measurable functions on a measure space (X, \mathfrak{B}, μ) converges almost everywhere to f on a set E with $\mu(E) < \infty$, then for an arbitrary $\varepsilon > 0$ we can find a set F ($F \subset E$, $F \in \mathfrak{B}$) such that $\mu(E - F) < \varepsilon$ and f_n converges uniformly on F . If $X = \mathbf{R}^n$ and \mathfrak{B} is either the class of Borel sets or the class of Lebesgue measurable sets, then the set F can be chosen to be a closed set (**Egorov's theorem**).

For a finite measurable function $f(x)$ defined on the real line, there exists a sequence $\{h_n\}$ such that $\lim_{n \rightarrow \infty} f(x + h_n) = f(x)$ almost everywhere (H. Auerbach).

The functional equation $f(x + y) = f(x) + f(y)$ has infinitely many nonmeasurable solutions (G. Hamel; — 388 Special Functional Equations).

Let \mathfrak{A}_i , $i = 1, 2$, be σ -algebras on X_i , $i = 1, 2$, respectively. A mapping $f: X_1 \rightarrow X_2$ is said to be measurable $\mathfrak{A}_1/\mathfrak{A}_2$ if $f^{-1}(A_2) \in \mathfrak{A}_1$ for every $A_2 \in \mathfrak{A}_2$. When \mathfrak{A}'_2 generates \mathfrak{A}_2 , $f: X_1 \rightarrow X_2$ is measurable $\mathfrak{A}_1/\mathfrak{A}_2$ if $f^{-1}(A'_2) \in \mathfrak{A}_1$ for every $A'_2 \in \mathfrak{A}'_2$. For example, a mapping $f: \mathbf{R}^1 \rightarrow \mathbf{R}^1$ is Borel (Lebesgue) measurable if and only if

f is measurable $\mathfrak{B}^1/\mathfrak{B}^1(\mathfrak{M}_1/\mathfrak{B}^1)$, where \mathfrak{B}^1 and \mathfrak{M}_1 denote the Borel subsets and the Lebesgue measurable subsets of \mathbf{R}^1 . Measurability of mappings is preserved by compositions, namely, if $f: X_1 \rightarrow X_2$ and $g: X_2 \rightarrow X_3$ are measurable $\mathfrak{A}_1/\mathfrak{A}_2$ and $\mathfrak{A}_2/\mathfrak{A}_3$, respectively, then $g \circ f: X_1 \rightarrow X_3$ is measurable $\mathfrak{A}_1/\mathfrak{A}_3$. If the projection $\pi_a: \prod_{i \in I} X_i \rightarrow X_a$ is measurable, then $\prod_{i \in I} \mathfrak{A}_i/\mathfrak{A}_a$. Since $\pi_i^{-1}(A_i)$ ($i \in I, A_i \in \mathfrak{A}_i$) generate $\prod_{i \in I} \mathfrak{A}_i$, a mapping $f: X \rightarrow \prod_{i \in I} X_i$ is measurable $\mathfrak{A}/\prod_{i \in I} \mathfrak{A}_i$ if $\pi_i \circ f: X \rightarrow X_i$ is measurable $\mathfrak{A}_i/\mathfrak{A}$ for every $i \in I$, where \mathfrak{A} is a σ -algebra on X .

K. Image Measures

Let μ be a regular measure on a measurable space (X, \mathfrak{B}_X) , i.e., the completion of a measure on \mathfrak{B}_X . Then every mapping $f: X \rightarrow Y$ measurable $\mathfrak{B}_X/\mathfrak{B}_Y$ induces a σ -algebra on Y :

$$\mathfrak{B} = \{B \subset Y \mid f^{-1}(B) \text{ is } \mu\text{-measurable}\},$$

and a measure on \mathfrak{B} :

$$\nu(B) = \mu(f^{-1}(B)).$$

The measure ν is called the **image measure** of μ under the mapping f ; this is denoted by μf^{-1} or $f_*\mu$. Obviously, $\nu(Y) = \mu(X)$, and ν is complete. Although \mathfrak{B} includes \mathfrak{B}_Y by virtue of the measurability of f , ν is not regular in general. However, ν is regular if μ is a bounded measure and if both (X, \mathfrak{B}_X) and (Y, \mathfrak{B}_Y) are analytic measurable spaces (\rightarrow 22 Analytic Sets I). Therefore the image measure of a regular probability measure under a measurable mapping is also a regular probability measure in practically all useful cases.

L. Related Topics

(i) **Integration.** For a nonnegative measurable function $f(x)$ on (X, \mathfrak{B}, m) , we can define the integral of $f(x)$ on a set $E \in \mathfrak{B}$, denoted by $\int_E f(x) \mu(dx)$, $\int_E f(x) d\mu(x)$, $\int_E f d\mu$, or $\int_E f$. For a real measurable function $f(x)$, the integral is defined to be equal to $\int_E f^+ d\mu - \int_E f^- d\mu$, if at least one of these integrals is finite, where f^+ and f^- are the positive and negative parts of f (\rightarrow 221 Integration Theory B).

(ii) **Fubini's theorem for measures.** Let (X, \mathfrak{B}, μ) be the complete product measure space of $(X_i, \mathfrak{B}_i, \mu_i)$ ($i = 1, 2$), where both μ_1 and μ_2 are σ -finite and complete. For $E \in \mathfrak{B}$ we define the sections $E(x_1)$ and $E(x_2)$ by

$$E(x_1) = \{x_2 \mid (x_1, x_2) \in E\},$$

$$E(x_2) = \{x_1 \mid (x_1, x_2) \in E\}.$$

Then we have $E(x_1) \in \mathfrak{B}_2$ for almost all $(\mu_1)x_1 \in$

X_1 , $E(x_2) \in \mathfrak{B}_1$ for almost all $(\mu_2)x_2 \in X_2$, and

$$\begin{aligned} \mu(E) &= \int_{X_1} \mu_2(E(x_1)) \mu_1(dx_1) \\ &= \int_{X_2} \mu_1(E(x_2)) \mu_2(dx_2) \end{aligned}$$

(\rightarrow 221 Integration Theory E).

(iii) **The Radon-Nikodym theorem and the Lebesgue decomposition theorem for measures.**

Let μ and ν be σ -finite measures on (X, \mathfrak{B}) . If $\mu(E) = 0$ implies $\nu(E) = 0$, then ν is said to be **absolutely continuous** with respect to μ , denoted by $\nu \ll \mu$. $\nu \ll \mu$ if and only if ν is expressible in the form $\nu(E) = \int_E f d\mu$, where f is a nonnegative \mathfrak{B} -measurable function (the **Radon-Nikodym theorem**). Such an f is uniquely determined up to μ -measure 0 and is called the **Radon-Nikodym derivative** of ν with respect to μ , denoted by $d\nu/d\mu$. The concept opposite to absolute continuity is singularity. ν is said to be **singular** with respect to μ if there exists an $E \in \mathfrak{B}$ such that $\nu(E^c) = \mu(E) = 0$. If μ and ν are two arbitrary σ -finite measures, then ν is expressible uniquely as the sum of an absolutely continuous measure ν_1 and a singular measure ν_2 with respect to μ (**Lebesgue decomposition theorem**) (\rightarrow 221 Integration Theory D, 380 Set Functions C).

(iv) **Invariant Measures.** Let (X, \mathfrak{B}) be a measurable space and G a group of Borel isomorphic mappings from (X, \mathfrak{B}) to itself. A measure μ on (X, \mathfrak{B}) is said to be **invariant** under G if $\mu(E) = \mu(g^{-1}(E))$ for every $E \in \mathfrak{B}$ and every $g \in G$. For example, the Lebesgue measure m on $(\mathbf{R}^n, \mathfrak{B}^n)$ is invariant under the group of congruent transformations (\rightarrow 225 Invariant Measures).

(v) **The Lebesgue-Stieltjes measure.** Let f be a right continuous monotone increasing function on \mathbf{R} . Then there exists a unique measure μ on $(\mathbf{R}, \mathfrak{B}^1)$ such that $\mu((a, b]) = f(b) - f(a)$ for $a < b$. This measure is called the **Lebesgue-Stieltjes measure** induced by f (\rightarrow 166 Functions of Bounded Variation B).

(vi) **Baire measurable functions and universally measurable functions.** Let $f(x)$ be a real-valued function defined on a topological space X . If for every open set O the inverse image $f^{-1}(O)$ has the \dagger Baire property (resp. is measurable with respect to the completion of any σ -infinite \dagger Borel measure on X), then f is said to be **Baire measurable** (resp. **universally measurable** or **absolutely measurable**). Universally measurable functions can be defined similarly on a measurable space.

(vii) **Disintegration.** Let (S, \mathfrak{S}) and (T, \mathfrak{T}) be standard measurable spaces, ν a σ -finite measure on (S, \mathfrak{S}) , and $p: S \rightarrow T$ a ν -measurable mapping. If the image measure $\mu = \nu f^{-1}$ is σ -finite, there exists a unique family of measures

$\{v_t, t \in T\}$ on (S, \mathfrak{S}) such that (1) $t \mapsto v_t(E)$ is universally measurable for every $E \in \mathfrak{S}$, (2) v_t is concentrated on $p^{-1}(t)$ for almost every t with respect to μ , and (3) the following equality holds:

$$v(E) = \int_T \mu(dt) v_t(E), \quad E \in \mathfrak{S}.$$

This expression is called the **disintegration** of v with respect to μ . Every v -measurable function f on S is v_t -measurable for almost every t with respect to μ , and the integral of f is expressible in the form of iterated integrals as follows:

$$\int_S f(s) v(ds) = \int_T \mu(dt) \int_S f(s) v_t(ds).$$

This corresponds to Fubini's theorem on product measures. When v is a probability measure on (S, \mathfrak{S}) , $p(s)$ is regarded as a (T, \mathfrak{T}) -valued random variable on (S, \mathfrak{S}, v) , and v_t is the conditional probability measure under the condition $p(s) = t$. (\rightarrow 342 Probability Theory E. For disintegration of measures on a topological space \rightarrow [3, ch. 9].)

References

- [1] H. Lebesgue, Intégral, longueur, aire, Ann. Mat. Pura Appl., (3) 7 (1902), 231–359.
- [2] S. Saks, Theory of the integral, Warsaw, 1937.
- [3] P. R. Halmos, Measure theory, Van Nostrand, 1950.
- [4] N. Bourbaki, Eléments de mathématique, Intégration Chap. 1–9, Actualités Sci. Ind., Hermann, 1965–1969.
- [5] H. L. Royden, Real analysis, Macmillan, second edition, 1963.
- [6] W. Rudin, Real and complex analysis, McGraw-Hill, second edition, 1974.
- [7] L. Schwartz, Probabilités cylindriques et applications radonifiantes, J. Fac. Sci. Univ. Tokyo, IA, 18 (1971), 139–286.

271 (XX.4) Mechanics

A. Newton's Three Laws of Motion

The study of laws governing the motion of bodies began with the laws of falling bodies, of which the first exact formulation was made by Galileo. But the general relationship between force and acceleration was first described by †Newton, who established **Newton's three laws of motion**; the mechanics based on them is **Newtonian mechanics**. Newton expounded

these laws in his famous book *Principia mathematica philosophiae naturalis* (1686–1687), where the law of gravitation and its application, problems of fluid motion, motions of the planets in the solar system, etc., were systematically treated.

Newton's first law. A body continues its state of rest or uniform motion in a straight line unless it is compelled to change that state by external action (i.e., force). This is also called the **law of inertia**.

Newton's second law. The rate of change of momentum is proportional to force and is in the direction in which the force acts. Here, **momentum** is defined as the product mv of the mass m and the velocity v . This law can be expressed as $d(mv)/dt = F$, where F is the force expressed in an appropriately chosen system of units. Since dv/dt is the acceleration a , the law takes the form $ma = F$, when the mass is constant. These equations are called **equations of motion**. The second law is often simply called the **law of motion**.

Newton's third law. When two bodies 1 and 2 in the same system interact, the force exerted on body 1 by body 2 is equal and opposite in direction to that exerted on body 2 by body 1. This law is called the **law of action and reaction**, or simply the **law of reaction**.

Various attempts at a rigorous axiomatization of Newtonian mechanics have been made, beginning with E. Mach's work. At the beginning of the 20th century, it was found that Newtonian mechanics requires modification when bodies travel at speeds approaching the speed of light or when it is applied to physical systems of molecular size or smaller. These modifications led to the establishment of the theory of †relativity and †quantum mechanics. In contrast to these later theories, Newtonian mechanics is called **classical mechanics**.

B. Newton's Law of Gravitation

Kepler discovered the following three laws for the motion of planets around the sun (valid within the accuracy in observation available at the time):

Kepler's first law. The orbit of a planet is an ellipse with the sun at one of its foci.

Kepler's second law. The area swept per unit time by the straight line segment joining the planet and the sun is independent of the position of the planet in its orbit.

Kepler's third law. The square of the period (the time needed for the planet to go around the orbit once) is proportional to the cube of the major axis of the orbit.

From these empirical laws, Newton deduced his **law of universal gravitation**: Between any

pair of point particles, with masses m_1 and m_2 and at a distance r , there arises an attractive force along the line joining the two points; the magnitude of this force is given by $F = Gm_1m_2r^{-2}$, where G is a universal constant (approximately 6.670×10^{-8} dyne $\text{cm}^2 \text{g}^{-2}$) called the gravitational constant.

C. Kinetic and Potential Energies

If the force \mathbf{F} on a particle is a function of the position \mathbf{x} of the particle and is the gradient $-\nabla U$ of a time-independent (scalar) function $U(\mathbf{x})$, called the **potential**, then

$$E = \frac{1}{2}mv^2 + U(\mathbf{x}) \quad \left(\mathbf{v} = \frac{d\mathbf{x}}{dt} \right)$$

is a constant of the motion of the particle.

E is called the **total energy**, while the first and second terms on the right-hand side are called the **kinetic energy** and **potential energy**, respectively.

For a system of n particles at points $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ with masses m_1, \dots, m_n , suppose that the force acting on the particle at $\mathbf{x}^{(j)}$ is $-\nabla^{(j)}U$ (the gradient of U relative to $\mathbf{x}^{(j)}$) for a common potential function $U(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$. Then

$$E = \sum_{j=1}^n \frac{1}{2}m_j v_j^2 + U(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \quad \left(\mathbf{v}_j = \frac{d\mathbf{x}^{(j)}}{dt} \right)$$

is the constant total energy of the motion. For example, Newton's gravitational force acting among a number of particles can be described by the **Newtonian potential**

$$U = \sum_{i < j} -Gm_i m_j |\mathbf{x}^{(i)} - \mathbf{x}^{(j)}|^{-1}.$$

A potential U which is a sum of functions depending only on a pair of coordinates as in the above example is called a **two-body interaction**.

D. Apparent Force

The coordinate system in which Newton's three laws of motion hold is called an **inertial system**. In some cases (e.g., on a rotating sphere such as the earth) it is more convenient to use a moving coordinate system, in which the equation of motion derived from Newton's second law by a coordinate transformation (from an inertial system to the moving coordinate system) takes a form similar to the second law except for an additional **apparent force** to be added to the force in the original equation of motion. For a coordinate system rotating at a constant angular velocity $\boldsymbol{\omega}$ (relative to an inertial system), the apparent force consists of the **centrifugal force**, of magnitude $m\omega^2\rho$ (ρ being the distance to the axis of rotation),

which pushes the particle away from the axis of rotation (along the perpendicular), and the **Coriolis force** $2m\mathbf{v} \times \boldsymbol{\omega}$ (\times denotes the vector product), which bends the motion of the particle in a direction perpendicular to both the axis of rotation and the velocity of the particle.

E. Dynamics of Rigid Bodies

A **rigid body** is defined as a system of particles whose mutual distances are permanently fixed. Like a point particle, it is an ideal concept introduced into mechanics to simplify the theoretical treatment. Actual solid bodies can in most cases be regarded as rigid under the action of forces of ordinary magnitudes. Since a rigid body can be imagined to be made up of an infinite number of particles, the equations of motion for a system of particles can also be applied to it. Thus the motion of a rigid body can be completely determined by the theorems of momentum and angular momentum.

The **momentum** of a rigid body is defined by

$$\mathbf{Q} = \int_K \frac{d\mathbf{r}}{dt} dm,$$

where dm is the mass of the volume element at a point \mathbf{r} of the rigid body K and $d\mathbf{r}/dt$ is its velocity. If the external forces acting on K are denoted by \mathbf{F}_i ($i = 1, 2, \dots$), we have

$$d\mathbf{Q}/dt = \sum \mathbf{F}_i,$$

which expresses the **theorem of momentum**.

If the velocity and acceleration of the **center of gravity** (**center of mass** or **barycenter**) $\int \mathbf{r} dm / \int dm$ of the rigid body are denoted by \mathbf{V}_G and \mathbf{A}_G , respectively, and the **mass** $\int dm$ by m , we have $\mathbf{Q} = m\mathbf{V}_G$, and the theorem of momentum becomes $m d\mathbf{V}_G/dt = m\mathbf{A}_G = \sum \mathbf{F}_i$.

The **angular momentum** of a rigid body K about an arbitrary point \mathbf{r}_0 is defined by

$$\mathbf{H} = \int_K (\mathbf{r} - \mathbf{r}_0) \times \frac{d\mathbf{r}}{dt} dm.$$

If \mathbf{P}_i is the vector from \mathbf{r}_0 to the point at which \mathbf{F}_i acts, we have

$$d\mathbf{H}/dt = \sum (\mathbf{P}_i \times \mathbf{F}_i) = \mathbf{G}.$$

This is called the **theorem of angular momentum**.

For the case of a rigid body with one point \mathbf{r}_0 fixed, the angular momentum \mathbf{H} and the angular velocity $\boldsymbol{\omega}$ are related by

$$H_x = A\omega_x - F\omega_y - E\omega_z,$$

$$H_y = -F\omega_x + B\omega_y - D\omega_z,$$

$$H_z = -E\omega_x - D\omega_y + C\omega_z,$$

where $H_x, H_y, H_z; \omega_x, \omega_y, \omega_z$ are the components of \mathbf{H} and $\boldsymbol{\omega}$ in the xyz -coordinate system

fixed in space with its origin at the fixed point \mathbf{r}_0 , and

$$A = \int (y^2 + z^2) dm, \quad B = \int (z^2 + x^2) dm,$$

$$C = \int (x^2 + y^2) dm, \quad D = \int yz dm,$$

$$E = \int zx dm, \quad F = \int xy dm,$$

with the integrals taken over the whole rigid body. We call A , B , and C the **moments of inertia** about the x -, y -, and z -axes, respectively, and D , E , F the corresponding **products of inertia**. The rotational motion of a rigid body with one axis fixed is completely determined by the theorem of angular momentum. However, for rotation about a fixed point, the theorem is not very convenient to use, because A , B , C , D , E , and F are generally unknown functions of time.

The \dagger quadric

$$Ax^2 + By^2 + Cz^2 + 2Dyz + 2Ezx + 2Fxy = 1$$

represents an ellipsoid with its center at the origin, called the **ellipsoid of inertia**. If the principal axes ξ , η , and ζ are taken as coordinate axes, the equation of the ellipsoid of inertia becomes $A\xi^2 + B\eta^2 + \Gamma\zeta^2 = 1$, where A , B , Γ are the moments of inertia about the ξ -, η -, ζ -axes and are called the **principal moments of inertia**, while the ξ -, η -, ζ -axes themselves are called the **principal axes of inertia**. If the components of the angular momentum \mathbf{H} and angular velocity in the direction of the principal axes of inertia are denoted by (H_1, H_2, H_3) and $(\omega_1, \omega_2, \omega_3)$, respectively, then $H_1 = A\omega_1$, $H_2 = B\omega_2$, $H_3 = \Gamma\omega_3$. Furthermore, if the ξ -, η -, ζ -components of the resultant moment of the external forces $\mathbf{G} = \sum(\mathbf{P}_i \times \mathbf{F}_i)$ are denoted by (G_1, G_2, G_3) , then $d\mathbf{H}/dt = \mathbf{G}$ becomes

$$A d\omega_1/dt = G_1 + (B - \Gamma)\omega_2\omega_3,$$

$$B d\omega_2/dt = G_2 + (\Gamma - A)\omega_3\omega_1,$$

$$\Gamma d\omega_3/dt = G_3 + (A - B)\omega_1\omega_2.$$

These are called **Euler's differential equations**.

The study of the motion of a rigid body can be reduced mostly to the study of the motion of its ellipsoid of inertia, since the latter is attached to the rigid body. The method of describing the motion of a rigid body by means of its ellipsoid of inertia is known as **Poinsot's representation**. The motions of two bodies having equal ellipsoids of inertia are the same if the external forces acting on them have equal resultant moments, even if their geometric forms are different.

When a rigid body moves under no constraint, the motion of its center of gravity G

can be determined by the use of the theorem of momentum. Also, the rotational motion about the center of gravity can be found from $d\mathbf{H}'/dt = \mathbf{G}'$, which is a modification of the theorem of angular momentum. Here \mathbf{H}' and \mathbf{G}' are respectively the angular momentum and moment of external forces about the center of gravity. In this case also, simplification can be achieved by considering the equation in the reference system that coincides with the principal axes of the ellipsoid of inertia with its center at the center of gravity.

F. Analytical Dynamics

Mechanics as originally formulated by Newton was geometric in nature, but later L. Euler, J. L. Lagrange, and others developed the analytical method of treating mechanics that is now called **analytical dynamics**. Lagrange introduced **generalized coordinates** q_j ($j = 1, 2, \dots, f$, where f is the number of **degrees of freedom** of the system considered), which uniquely represent the configuration of the dynamical system, and derived **Lagrange's equations of motion**:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) - \frac{\partial \mathcal{L}}{\partial q_j} = 0, \quad j = 1, 2, \dots, f,$$

where $\dot{q}_j = dq_j/dt$, and $\mathcal{L} = T - U$ (T = kinetic energy, U = potential energy) is a function of q_j and \dot{q}_j called the **Lagrangian function**. Later, W. R. Hamilton introduced

$$p_j = \partial T / \partial \dot{q}_j,$$

$$H = \sum p_j \dot{q}_j - \mathcal{L} = H(p_1, \dots, p_f; q_1, \dots, q_f)$$

and transformed the equations to **Hamilton's canonical equations**:

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}, \quad j = 1, 2, \dots, f.$$

Here p_j is the **generalized momentum** conjugate to q_j , and q_j , p_j are called **canonical variables**.

If the functions representing the configuration of the dynamical system in terms of q_j do not explicitly contain the time t , the **Hamiltonian function** (or **Hamiltonian**) H coincides with the total energy of the system $T + U$.

The transformation $(p, q) \rightarrow (P, Q)$ under which canonical equations preserve their form is called a **canonical transformation**. It is given by

$$p_j = \frac{\partial W}{\partial q_j}, \quad P_j = -\frac{\partial W}{\partial Q_j}, \quad K = H + \frac{\partial W}{\partial t},$$

where $W = W(q_1, \dots, q_f; Q_1, \dots, Q_f)$ and K is the Hamiltonian of the transformed system. The set of canonical transformations forms a group, called a **group of canonical transforma-**

tions. An infinitesimal transformation is given by

$$dp_j = -\varepsilon \frac{\partial S}{\partial q_j}, \quad dq_j = \varepsilon \frac{\partial S}{\partial p_j},$$

where ε is an infinitesimal constant. Here S is an arbitrary function and is said to be the **generating function** of the infinitesimal transformation. Canonical equations can be interpreted to mean that the variations of p and q during the time interval $\varepsilon = dt$ are the infinitesimal canonical transformations whose generating function is $H(p, q, t)$.

The variation of an arbitrary function $F(p, q)$ under an infinitesimal transformation is given by

$$dF = \varepsilon(F, S),$$

where

$$(u, v) = \sum_j \left(\frac{\partial u}{\partial q_j} \frac{\partial v}{\partial p_j} - \frac{\partial u}{\partial p_j} \frac{\partial v}{\partial q_j} \right) = \sum_j \frac{\partial(u, v)}{\partial(q_j, p_j)}$$

is **Poisson's bracket**. Therefore the time rate of change of a dynamical quantity $F(p, q)$ can be written as

$$dF/dt = (F, H).$$

Thus the function $F(p, q)$ that satisfies $(F, H) = 0$ is an integral of the canonical equations.

If a canonical transformation $(p, q) \rightarrow (P, Q)$ such that $P_j = \alpha_j$, $Q_j = \beta_j$ are constant is found, the motion of the system can be determined by

$$p_j = \frac{\partial W}{\partial q_j}, \quad \beta_j = \frac{\partial W}{\partial \alpha_j},$$

where W is the complete solution of the **Hamilton-Jacobi differential equation**:

$$\frac{\partial W}{\partial t} + H\left(\frac{\partial W}{\partial q_1}, \dots, \frac{\partial W}{\partial q_f}; q_1, \dots, q_f, t\right) = 0$$

(\rightarrow 82 Contact Transformations).

G. Theory of Elasticity

(1) General remarks. Suppose that a solid body is deformed elastically by the action of external forces. We may inquire about the magnitudes of deformation, stress, and strain caused by the external forces at each point of the body. The **theory of elasticity** studies this problem, assuming the body to be a continuum and utilizing classical mechanics as a basis, and "endeavours to obtain results which shall be practically important in applications to architecture, engineering and all other useful arts in which the material of construction is solid" [1].

Cartesian coordinates (x, y, z) are employed for defining the 3-dimensional space contain-

ing the body. We confine ourselves to the small-displacement theory of elasticity, in which the components of the displacement $\mathbf{u} = (u, v, w)$ of an arbitrary point $P(x, y, z)$ of the body are assumed to be small enough to justify the linearization of the governing differential equations and boundary conditions.

(2) Stress. Consider an infinitesimal rectangular parallelepiped enclosed by the following six surfaces:

$$x = \text{const}, \quad y = \text{const}, \quad z = \text{const},$$

$$x + dx = \text{const}, \quad y + dy = \text{const}, \quad z + dz = \text{const},$$

in the body. The **stress** at the point $P(x, y, z)$ is defined as those internal forces per unit area acting on the six surfaces of the parallelepiped. It has nine components, which are usually represented by

$$\begin{pmatrix} \sigma_x & \tau_{yx} & \tau_{zx} \\ \tau_{xy} & \sigma_y & \tau_{zy} \\ \tau_{xz} & \tau_{yz} & \sigma_z \end{pmatrix},$$

where the σ and τ are called **normal** and **shearing** (or **tangential**) **stresses**, respectively. These nine quantities form a tensor called the **stress tensor**. Since it can be shown that $\tau_{xy} = \tau_{yx}$, $\tau_{yz} = \tau_{zy}$, $\tau_{zx} = \tau_{xz}$, the stress tensor is a symmetric tensor. By considering equilibrium conditions of the parallelepiped, the equations of equilibrium are found to be

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z} + \bar{X} = 0, \dots, \dots,$$

where \bar{X}, \dots are **body forces** per unit volume.

(3) Strain. The infinitesimal rectangular parallelepiped fixed to the body at the point P before deformation is transformed, after deformation, into an infinitesimal parallelepiped which is no longer rectangular. The **strain** at the point P is defined as the changes caused by the deformation of the parallelepiped: extensions of the three sides and changes from right angles of the three angles formed by the three sides of the parallelepiped. Thus the strain has six components, usually represented by

$$(\varepsilon_x, \varepsilon_y, \varepsilon_z, \gamma_{yz}, \gamma_{zx}, \gamma_{xy}),$$

where ε and γ are called **elongation** and **shearing strains**, respectively. These six quantities, with slight modification, form a symmetric tensor called the **strain tensor**.

(4) Strain-displacement relations. In small-displacement theory, the strain-displacement relations are given in linear form by

$$\varepsilon_x = \frac{\partial u}{\partial x}, \quad \varepsilon_y = \frac{\partial v}{\partial y}, \quad \dots, \quad \gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}.$$

(5) Stress-strain relations. In the theory of elasticity, stress and strain are assumed to

obey a linear relationship called **Hooke's law**:

$$\{\sigma\} = [A] \{\varepsilon\},$$

where

$$\{\sigma\}^T = [\sigma_x, \sigma_y, \sigma_z, \tau_{yz}, \tau_{zx}, \tau_{xy}],$$

$$\{\varepsilon\}^T = [\varepsilon_x, \varepsilon_y, \varepsilon_z, \gamma_{yz}, \gamma_{zx}, \gamma_{xy}],$$

and $[A]$ is a symmetric positive definite matrix. For an isotropic body, the relations are

$$E\varepsilon_x = \sigma_x - \nu(\sigma_y + \sigma_z), \quad \dots, \quad G\gamma_{xy} = \tau_{xy},$$

where E , ν , and $G = E/2(1 + \nu)$ are elastic constants called the **modulus of elasticity in tension** or **Young's modulus**, **Poisson's ratio**, and the **modulus of elasticity in shear** or the **modulus of rigidity**, respectively.

(6) **Boundary conditions.** The surface of the body can be divided into two parts with regard to boundary conditions: the part S_1 over which the boundary conditions are prescribed in terms of external forces and the part S_2 over which the boundary conditions are prescribed in terms of displacements. Obviously $\partial V = S_1 + S_2$, where ∂V is the whole surface of the body.

(7) **Small-displacement theory of elasticity.** We have seen that the equations which govern the problem are 3 equations of equilibrium, 6 strain-displacement relations, and 6 stress-strain relations in terms of 15 unknowns, namely, 6 stress components, 6 strain components, and 3 displacement components. Thus our problem is reduced to a boundary value problem in which these 15 field equations are to be solved under the specified boundary conditions. Since all the field equations and boundary conditions are linear with respect to the unknowns under the assumption that the displacements are small, we obtain linear relationships between the external load and resulting deformation of the body; this is the **small-displacement theory of elasticity**. It should be remembered, however, that the assumption of small displacement sets limits to the application of the theory to practical problems.

(8) **Variational principles.** Several **variational principles** have been formulated in the small-displacement theory of elasticity. These include the principle of minimum potential energy $[u]$, the generalized principle $[\sigma, \varepsilon, u]$, the Hellinger-Reissner principle $[\sigma, u]$, the principle of minimum complementary energy $[\sigma]$, and so forth, where the symbols in the brackets represent independent functions subject to variation. In connection with the aforementioned variational principles, **variational principles with relaxed continuity requirements** have also been formulated by re-

laxing the continuity requirements imposed on admissible functions.

One of the practical advantages of these variational principles is that they often provide the problem with approximate formulations and approximate methods of solution, among which the **Rayleigh-Ritz method** is well known. Theories of beams, plates, shells, and multi-component structures are typical examples of such approximate formulations. Recently, these variational principles have been found to provide effective bases for the formulations of the ^{*}finite element method.

(9) **Notation.** Various symbols are used for stress and strain. For example, σ , τ and ε , γ are commonly used in the engineering literature. However, in Love's treatise [6], X_x , Y_x , Z_x and e_{xx} , e_{xy} , e_{xz} are used in place of σ_x , τ_{xy} , τ_{xz} and ε_x , γ_{xy} , γ_{xz} , respectively. Also, various notations are used for elastic constants. E is widely used, while G and ν are less common. Love uses μ and σ for G and ν , respectively. In the engineering literature the reciprocal of Poisson's ratio $m (= 1/\nu)$ is used and is called the **Poisson number**.

(10) **Finite-displacement theory of elasticity.** When the displacement of the body is no longer small (infinitesimal) but is finite, we should abandon small-displacement theory and instead employ **finite-displacement theory**, in which stress and strain are defined in a manner different from that in small-displacement theory, keeping in mind the difference between spatial and material variables. Thus equations of equilibrium, strain-displacement relations, and the boundary conditions on S_1 become nonlinear equations with stress and displacement components as unknowns, although the stress-strain relations remain linear. Thus the problem is reduced to solving a nonlinear boundary value problem, sometimes called a nonlinear elasticity problem. Variational principles have been formulated for finite-displacement theory and are frequently used in the formulation of approximate methods of solution.

When the stress becomes large enough to exceed the so-called **elastic limit**, where the linear stress-strain relationships cease to hold, the theory of elasticity is no longer valid and should be replaced by the **theory of plasticity** [9, 10].

References

- [1] E. T. Whittaker, A treatise on the analytical dynamics of particles and rigid bodies, Cambridge Univ. Press, fourth edition, 1927.

Meromorphic Functions

- [2] A. Sommerfeld, *Mechanics*, translated by M. O. Stern, Academic Press, 1952. (Original in German, 1943.)
- [3] J. B. Marion, *Classical dynamics of particles and systems*, Academic Press, 1965.
- [4] V. I. Arnold, *Mathematical methods of classical mechanics*, translated by K. Vogtmann and A. Weinstein, Springer, 1978. (Original in Russian, 1974.)
- [5] R. Abraham and J. E. Marsden, *Foundations of mechanics*, Benjamin/Cummings, second edition, 1978.
- [6] A. E. H. Love, *A treatise on the mathematical theory of elasticity*, Cambridge Univ. Press, fourth edition, 1934.
- [7] S. Timoshenko and J. N. Goodier, *Theory of elasticity*, McGraw-Hill, 1951.
- [8] A. E. Green and W. Zerna, *Theoretical elasticity*, Oxford Univ. Press, 1954.
- [9] R. Hill, *Mathematical theory of plasticity*, Oxford Univ. Press, 1950.
- [10] K. Washizu, *Variational methods in elasticity and plasticity*, Pergamon, second edition, 1975.

272 (XI.7)

Meromorphic Functions

A. General Remarks

A single-valued [†]analytic function in a domain D in the complex plane \mathbb{C} is called **meromorphic** in D if it has no singularities other than [†]poles. A function that is meromorphic in the whole complex plane including the point at infinity is a rational function (**Liouville's theorem**). Specifically, if a function is meromorphic in the domain \mathbb{C} , then the function is called simply a **meromorphic function**, and a meromorphic function that is not a rational function is called a **transcendental meromorphic function**. A meromorphic function $f(z)$ can be represented as a quotient of two [†]entire functions. Let $\{z_k\}$ ($k = 1, 2, \dots$) be poles of $f(z)$, and let $f_k(z) = a_{n_k}^{(k)}/(z - z_k)^{n_k} + \dots + a_1^{(k)}/(z - z_k)$ denote the [†]singular parts of $f(z)$ at z_k ($k = 1, 2, \dots$). Then $f(z)$ can also be written in the form

$$f(z) = g(z) + \sum_{k=1}^{\infty} (f_k(z) - p_k(z)),$$

where $g(z)$ is an entire function and the $p_k(z)$ ($k = 1, 2, \dots$) are rational entire functions (**Weierstrass's theorem**). Assume that a sequence $\{z_k\}$ ($k = 1, 2, \dots$) converges only to the point at infinity and that $f_k(1/(z - z_k))$ ($k = 1, 2, \dots$) are rational entire functions of $1/(z - z_k)$ which have no constant terms. Then

there exists a meromorphic function of z with $f_k(1/(z - z_k))$ as its [†]singular part at z_k (**Mittag-Leffler's theorem**).

B. Nevanlinna Theory

The theory of meromorphic functions can be considered an extension of the theory of entire functions. In particular, value distribution theory, originating in Picard's theorem, was studied by many people, and in 1925 R. Nevanlinna published a systematic theory unifying the results obtained until then. This is called **Nevanlinna theory**.

We let $f(z)$ denote a meromorphic function in $|z| < R \leq +\infty$, and when we say that $f(z)$ takes on a value, the value may be ∞ . For a value α , $n(r, \alpha)$ denotes the number of **α -points** of $f(z)$, i.e., points z with $f(z) = \alpha$, in $|z| \leq r < R$, where each α -point is counted with its multiplicity. We set

$$N(r, \alpha) = \int_0^r \frac{n(t, \alpha) - n(0, \alpha)}{t} dt + n(0, \alpha) \log r,$$

$$m(r, \alpha) = \frac{1}{2\pi} \int_0^{2\pi} \log^+ \left| \frac{1}{f(re^{i\theta}) - \alpha} \right| d\theta$$

if $\alpha \neq \infty$, and

$$N(r, \infty) = \int_0^r \frac{n(t, \infty) - n(0, \infty)}{t} dt + n(0, \infty) \log r,$$

$$m(r, \infty) = \frac{1}{2\pi} \int_0^{2\pi} \log^+ |f(re^{i\theta})| d\theta$$

if $\alpha = \infty$, where $\log^+ a = \max(\log a, 0)$ for $a > 0$.

The functions N and m are called the **counting function** and **proximity function** of $f(z)$, respectively, and $T(r) = T(r, f) = m(r, \infty) + N(r, \infty)$ is the **order function** (or **characteristic function**) of $f(z)$. $T(r)$ is an increasing function of r and a [†]convex function of $\log r$ and is useful for expressing $f(z)$ as an infinite product, etc.

The following relation holds among $T(r)$, $m(r, \alpha)$, and $N(r, \alpha)$ for any α :

$$T(r) = m(r, \alpha) + N(r, \alpha) + O(1), \quad (1)$$

where $O(1)$ is [†]Landau's symbol (**Nevanlinna's first fundamental theorem**). By this theorem, if a bounded remainder is disregarded, then $m(r, \alpha) + N(r, \alpha)$ is equal to $T(r)$ for all α . This equality thus demonstrates a beautifully balanced distribution of α -points.

We see that $N(r, \alpha)$ is in a sense the mean value of the number of α -points in $|z| \leq r$, and $m(r, \alpha)$ is the mean proximity to α of the value $f(z)$ on $|z| = r$. If the term \log^+ in the definition of the proximity function is replaced by the logarithm of the reciprocal of the chordal distance between $f(re^{i\theta})$ and α on the complex sphere, then the remainder term in (1) is

eliminated. Hence the definition of the proximity function is sometimes given in this form.

C. The Order of Meromorphic Functions

For an entire function $f(z)$, the equality

$$\limsup_{r \rightarrow \infty} \frac{\log T(r)}{\log r} = \limsup_{r \rightarrow \infty} \frac{\log \log M(r)}{\log r}$$

holds, where $M(r) = M(r, f) = \max_{|z|=r} |f(z)|$. Since the right-hand side is the order of $f(z)$ (\rightarrow 429 Transcendental Entire Functions), we define the **order (lower order)** ρ of a meromorphic function $f(z)$ by

$$\rho = \limsup_{r \rightarrow \infty} \frac{\log T(r)}{\log r} \quad \left(\liminf_{r \rightarrow \infty} \frac{\log T(r)}{\log r} \right).$$

The **order** of a meromorphic function in $|z| < R$ is also defined by

$$\rho = \limsup_{r \rightarrow R} \frac{\log T(r)}{\log(1/(R-r))}.$$

D. Meromorphic Functions on a Disk

The order function $T(r)$ is bounded if and only if $f(z)$ can be represented as the quotient of two bounded holomorphic functions $h_1(z), h_2(z)$ in $|z| < R$ (Nevanlinna). If $T(r)$ is bounded, $\lim_{r \rightarrow R} f(re^{i\theta})$ exists and is finite for every $\theta, 0 \leq \theta < 2\pi$, except possibly for a set with linear measure zero (P. Fatou and Nevanlinna). Among functions $f(z)$ such that $\lim_{r \rightarrow R} T(r) = \infty$, those satisfying

$$\limsup_{r \rightarrow R} \frac{T(r)}{\log(1/(R-r))} = \infty$$

have properties similar to those of transcendental meromorphic functions.

E. Meromorphic Functions in the Whole Finite Plane

Any meromorphic function such that

$$\limsup_{r \rightarrow R} \frac{T(r)}{\log r} < K$$

is a rational function. If $f(z)$ is a meromorphic function of order ρ and $\{r_j(\alpha)\}$, $r_j(\alpha) \leq r_{j+1}(\alpha)$ ($j = 1, 2, \dots$) is the set of absolute values of α -points, then $\sum_{j=1}^{\infty} (1/r_j(\alpha))^{\rho+\epsilon}$ converges for any α . Furthermore,

$$f(z) = z^k e^{P(z)}$$

$$\times \lim_{r \rightarrow \infty} \prod_{\substack{|a_\mu| \leq r \\ |b_\nu| \leq r}} \left(1 - \frac{z}{a_\mu} \right) \exp \left(\frac{z}{a_\mu} + \dots + \frac{z^p}{p a_\mu^p} \right) \cdot \left(1 - \frac{z}{b_\nu} \right) \exp \left(\frac{z}{b_\nu} + \dots + \frac{z^p}{p b_\nu^p} \right),$$

Meromorphic Functions

where p is the smallest integer satisfying $p+1 > \rho$, the a_μ and b_ν are the zeros and poles of $f(z)$, respectively, k is an integer, and $P(z)$ is a polynomial of degree at most p (**Hadamard's theorem**).

Let $\alpha_1, \dots, \alpha_q$ ($q \geq 3$) be distinct values. Then for any meromorphic function in $|z| < R \leq \infty$,

$$(q-2)T(r) < \sum_{j=1}^q N(r, \alpha_j) - N_1(r) + D(r),$$

$$0 \leq r < R.$$

Here

$$N_1(r) = \int_0^r \frac{n_1(t) - n_1(0)}{t} dt + n_1(0) \log r,$$

$n_1(r)$ is the number of multiple points in $|z| \leq r$ (a multiple point of order k is counted $k-1$ times), and $D(r)$ is the remainder such that if $R = \infty$, then $D(r) < K(\log T(r) + \log r)$ for some K except possibly for values of r belonging to the union of a countable number of intervals with finite total length, and if $R < \infty$, then $D(r) < K(\log T(r) + \log(1/(R-r)))$ except possibly for the union of a countable number of intervals $\{I_j\}$ with $\sum_j \int_{I_j} d(1/(R-r)) < \infty$ (**Nevanlinna's second fundamental theorem**).

Several theorems on value distribution of meromorphic functions can be obtained directly from this theorem. For instance, if $f(z)$ is a transcendental meromorphic function, the equation $f(z) = \alpha$ has an infinite number of roots for every value α except for at most two values called **Picard's exceptional values** (**Picard's theorem**). For a meromorphic function of order ρ , $\lim_{r \rightarrow \infty} \sum_{r_j \leq r} (r_j(\alpha))^{-\lambda}$ ($\lambda < \rho$) diverges for every value α except for at most two values (**Borel's theorem**). A value α for which the series converges is called a **Borel exceptional value**. We call $\delta(\alpha) = \delta(\alpha, f) = 1 - \limsup_{r \rightarrow \infty} N(r, \alpha)/T(r)$ the **defect** of f . It always satisfies $0 \leq \delta(\alpha) \leq 1$, and the values with $\delta(\alpha) > 0$ are called **Nevanlinna's exceptional values**. The number of values α (may be ∞) with $\delta(\alpha) > 0$ is at most countable for any meromorphic function $f(z)$, and $\sum_{i=1}^{\infty} \delta(\alpha_i) \leq 2$. There are many studies concerning the values α with $\delta(\alpha) = 0$.

F. Julia Directions

Among functions that have an essential singularity at the point at infinity and are meromorphic in the whole plane, there are some that possess no Julia directions. These functions, called **Julia exceptional functions**, are of order 0. A necessary and sufficient condition for $f(z)$ to be a Julia exceptional function is that $f(z)$ can be written in the form $z^m \prod_\mu (1 - z/a_\mu) / \prod_\nu (1 - z/b_\nu)$ (A. Ostrowski,

Meromorphic Functions

1925). Concerning zeros a_μ and poles b_ν of $f(z)$, the following three properties are obtained using the theory of normal families due to P. Montel: (1) There are constants K_1 , K_2 , and K_3 independent of r such that $|n(r, \infty) - n(r, 0)| < K_1$, $n(2r, \infty) - n(r, \infty) < K_2$, and $n(2r, 0) - n(r, 0) < K_3$. (2) There are constants K_4 and K_5 such that for any p and q ,

$$|a_p|^m \prod_{|a_\mu| < |a_p|} \left| \frac{a_p}{a_\mu} \right| \bigg/ \prod_{|b_\nu| < |a_p|} \left| \frac{a_p}{b_\nu} \right| < K_4,$$

$$|b_q|^{-m} \prod_{|b_\nu| < |b_q|} \left| \frac{b_q}{b_\nu} \right| \bigg/ \prod_{|a_\mu| < |b_q|} \left| \frac{b_q}{a_\mu} \right| < K_5.$$

(3) There exists an $\varepsilon > 0$ satisfying $|a_p/b_q - 1| \geq \varepsilon > 0$ for any p and q .

G. Valiron gave a precise form of Julia directions that corresponds to Borel's theorem (*Acta Math.*, 52 (1928), [3]). Namely, if the order ρ of a meromorphic function $f(z)$ in $|z| < \infty$ is positive and finite, then there exists a direction J defined by $\arg z = \alpha$ such that the zeros $z_\nu(a, \Delta)$ of $f(z) - a$ in any angular domain $\Delta: |\arg z - \alpha| < \delta$ containing J have the property $\sum_\nu |z_\nu(a, \Delta)|^{-(\rho-\varepsilon)} = \infty$ for any $\varepsilon > 0$ except for at most two values of a . The direction J is called a **Borel direction**.

G. Relations between Two or More Meromorphic Functions

Borel's unicity theorem can be stated as follows: Let f_j ($j = 1, \dots, n$) be nonvanishing entire functions satisfying $\sum_{j=1}^n f_j = 1$; then for some $(c_1, \dots, c_n) \neq (0, \dots, 0)$, $\sum_{j=1}^n c_j f_j = 0$. This is contained in the following theorem: Let f_j ($j = 1, \dots, n$) be transcendental entire functions such that $\sum_{j=1}^n f_j = 1$; then $\sum_{j=1}^n \delta(0, f_j) \leq n - 1$. If two meromorphic functions $f_1(z)$, $f_2(z)$ have the same α_j -points for five distinct values α_j ($j = 1, \dots, 5$) (where multiplicity is not taken into account), then they coincide everywhere. If the \dagger Riemann surface of the \dagger algebraic function $w(z)$ defined by a polynomial $P(z, w) = 0$ of z , w is of \dagger genus > 1 , it is impossible to find meromorphic functions $z = f(\zeta)$, $w = g(\zeta)$ that satisfy $P(f(\zeta), g(\zeta)) = 0$ (\dagger uniformization by meromorphic functions).

H. Asymptotic Values

If a meromorphic function $f(z) \rightarrow \alpha$ as $z \rightarrow \infty$ along a curve C , the value α and the curve C are called an **asymptotic value** and **asymptotic path**, respectively. Each (Picard's) exceptional value of $f(z)$ is an asymptotic value. For meromorphic functions, no simple relation is known between their order and the number of their asymptotic values. There exists a meromorphic function of order 0 with an infinite

number of asymptotic finite values. Some results analogous to those for entire functions are obtained for meromorphic functions by applying the theory of normal families. F. Marty established a systematic theory of normal families of meromorphic functions by using spherical distance.

I. Inverse Functions

Generally, the inverse function of a meromorphic function $w = f(z)$ is infinitely multiple-valued. Let $P(w, w_0)$ be a function element of the inverse function with center at w_0 , and let C be an arbitrary curve starting at w_0 and with ω its terminal point. For any domain S containing C , $P(w, w_0)$ can be continued analytically in S up to a point arbitrarily near ω (**Iversen's theorem**). Continue the function element analytically along each half-line starting at its center. Then the set of arguments of half-lines along which the analytic continuation meets a singularity at a finite point is of zero linear measure (**Gross's theorem**).

By considering the inverse image of the suitably cut Riemann surface of the inverse function, the z -plane can be divided into fundamental domains such that each domain is the inverse image of the whole w -plane (with suitable slits removed) and has a boundary each point of which is \dagger accessible from the inside of the domain, and the boundary curves of fundamental domains cluster nowhere in the plane.

For a meromorphic function $f(z)$, the set of functions $z' = \varphi(z)$ defined by $f(z') = f(z)$ (i.e., transformations between points that give $f(z)$ the same value) has the property of a \dagger hypergroup. If $\varphi(z)$ is single-valued, then it is a linear entire function, and if it is finitely multiple-valued, then it is an algebraic function. The \dagger cluster set of the inverse function at a transcendental singularity consists of only one point, ∞ , that is, it is an \dagger ordinary singularity. To an analytic continuation along a curve that determines a transcendental singularity of the inverse function there corresponds a curve in the z -plane terminating at ∞ . This curve is an asymptotic path of $f(z)$. Namely, the value $f(z)$ tends to the coordinate of the transcendental singularity as $z \rightarrow \infty$ along this path. Each asymptotic value of a transcendental meromorphic function $w = f(z)$ corresponds to a transcendental singularity of its inverse function $z = \varphi(w)$, and if we consider two asymptotic paths to be the same if they correspond to the same singularity, then there exists a one-to-one correspondence between the set of asymptotic paths and the set of transcendental singularities of the inverse

function. The inverse function of any meromorphic function of order ρ has at most 2ρ [†]direct transcendental singularities if $\rho \geq 1/2$ and at most 1 such singularity if $\rho < 1/2$ (L. V. Ahlfors).

J. Theory of Covering Surfaces

Ahlfors established the theory of covering surfaces by a metricotopological method and in applying it, obtained Nevanlinna theory and many other results on meromorphic functions.

Let F_r denote the covering surface of the Riemann sphere F_0 with radius $1/2$; F_r is the image of $|z| \leq r$ under a meromorphic function $w = f(z)$. The area of F_r divided by π , where π is the area of F_0 , is given by

$$A(r) = \frac{1}{\pi} \iint_{|z| \leq r} \frac{|f'(z)|^2}{(1 + |f(z)|^2)^2} \rho \, d\rho \, d\theta,$$
$$z = \rho e^{i\theta},$$

and is called the **mean number of sheets** of F_r . The length of the boundary of F_r is given by

$$L(r) = \int_{|z|=r} \frac{|f'(z)|}{1 + |f(z)|^2} |dz|.$$

The relation

$$T(r) = \int \frac{A(r)}{r} dr + O(1)$$

holds (T. Shimizu, Ahlfors).

Consider the Riemann surface of the inverse function of a meromorphic function $w = f(z)$ in $|z| < R \leq +\infty$. It has a countable number of components Q_v over a domain on the w -plane. Let Δ_v denote the inverse image of Q_v on the z -plane. If Δ_v together with its boundary is contained in $|z| < R$, the component Q_v is called an **island**, and otherwise, a **peninsula**.

Let D be a simply connected domain of the w -plane, $n(r, D)$ be the sum of the sheet numbers of the islands of F_r over D , and $m(r, D)$ be the sum of the areas of the peninsulas of F_r over D divided by the area of D . Then

$$m(r, D) + n(r, D) = A(r) + O(L(r)).$$

Let D_j ($j = 1, \dots, q$) ($q \geq 3$) be disjoint simply connected domains on the w -plane. Then

$$\sum_{j=1}^q n(r, D_j) - \sum_{j=1}^q n_1(r, D_j) > (q-2)A(r) - O(L(r)),$$

where $n_1(D)$ is the sum of the orders of branch points in all islands of F_r over D .

For a meromorphic function $f(z)$, $L(r) < A(r)^{1/2+\epsilon}$, where r satisfies $0 \leq r < \infty$ except for $r \in \bigcup_j I_j$ for some intervals I_j . Hence in the case where D_j is a point α_j , this inequality

yields

$$\sum_{j=1}^q n(r, \alpha_j) - \sum_{j=1}^q n_1(r, \alpha_j) > (q-2)A(r) - O(A(r)^{1/2+\epsilon})$$

with some exceptional intervals of values of r . This latter important inequality corresponds to Nevanlinna's second fundamental theorem. Let D_j ($j = 1, \dots, q$) ($q \geq 3$) be disjoint simply connected domains on the w -plane. If every simply connected island over D_j has at least μ_j sheets, then $\sum_{j=1}^q (1 - (1/\mu_j)) \leq 2$ (**disk theorem**). It follows from this theorem that given three disjoint disks D_j on the Riemann sphere, there is at least one D_j that has an infinite number of islands over it, and also that given five D_j , there exists at least one D_j that has a 1-sheeted island over it (**Ahlfors's five-disk theorem**). This theorem corresponds to [†]Bloch's theorem for entire functions. These theorems can also be obtained for meromorphic functions on a disk. Ahlfors established a more important theory by introducing a differential metric.

K. Recent Development

The Nevanlinna brothers raised several important problems, which gave strong motivation for later investigations. The first major breakthrough after World War II was given by A. Gol'dberg in 1956. He gave an example which has infinitely many deficient values (Nevanlinna's exceptional values). W. Hayman proved that $\sum \delta(a, f)^{\alpha}$ converges for $\alpha > 1/3$, and there are examples of meromorphic functions for which the series diverges for $\alpha < 1/3$. Finally, A. Weitsman showed that the series converges for $\alpha = 1/3$. The second major breakthrough was given by A. Edrei and W. Fuchs in 1959, whose works concern the following Nevanlinna theorem: Let $K(f)$ be

$$\limsup_{r \rightarrow \infty} \frac{N(r, 0) + N(r, \infty)}{T(r, f)}$$

and $\kappa(\rho) = \inf K(f)$, where \inf is taken over all meromorphic function f of order ρ . Then $\kappa(\rho) > 0$ if ρ is neither a positive integer nor ∞ . Furthermore, Nevanlinna gave a conjecture for an exact value of $\kappa(\rho)$. This conjecture is still open, although several estimates have appeared. In the case of entire functions having only negative zeros the conjecture was positively solved by S. Hellerstein and J. Williamson. Edrei and Fuchs proved that $\kappa(\rho) = 1$ for $0 \leq \rho \leq 1/2$ and $\kappa(\rho) = \sin \pi \rho$ for $1/2 < \rho \leq 1$. They also proved the ellipse theorem: Let $f(z)$ be a transcendental meromorphic function of order ρ ($0 \leq \rho \leq 1$). Put $u = 1 - \delta(a, f)$, $v = 1 - \delta(b, f)$. Then $u, v \in [0, 1]$ and $u^2 - 2uv \cos \pi \rho + v^2 \geq \sin^2 \pi \rho$. If further $u < \cos \pi \rho$, then $v = 1$.

For the sum of deficiencies Edrei proved that

$$\sum \delta(a, f) \leq \begin{cases} 1 - \cos \pi \rho & (0 \leq \rho \leq 1/2) \\ 2 - \sin \pi \rho & (1/2 < \rho \leq 1) \end{cases}$$

unless the number of deficient values is one. It was conjectured that $\sum \delta(a, f) = 2$ implies that the order ρ of f is a half integer, the number v of deficient values is at most 2ρ , and the value of $\delta(a, f)$ is a multiple of $1/\rho$. For the case of entire functions this conjecture is true (A. Pfluger). Weitsman proved that $v \leq 2\rho$. Other cases remain open. All the above results still hold even if the order is replaced by the lower order. The inverse problem was completely solved by D. Drasin. Many of the above results depend on the concept of Pólya peaks.

N. V. Govorov and V. Petrenko proved independently that

$$\liminf_{r \rightarrow \infty} \frac{\log M(r, f)}{T(r, f)} \leq \pi \rho \quad \text{for } \rho > 1/2$$

for every entire function of order ρ . This was a conjecture made by R. Paley. For $\rho \leq 1/2$, an exact upper bound $\pi \rho / \sin \pi \rho$ was given by Valiron. Furthermore, the following result was proved by Edrei and Fuchs:

$$\limsup_{r \rightarrow \infty} \frac{N(r, 0)}{\log M(r, f)} \geq \frac{\sin \pi \rho}{\pi \rho} \quad (0 \leq \rho < 1).$$

L. History

The value distribution theory of meromorphic functions had its inception with the classical Picard theorem. It first appeared as the value distribution theory of entire functions and was developed into a well-organized field by way of the Nevanlinna theory and the Ahlfors theory of covering surfaces. In recent years, emphasis has also been placed on the study of meromorphic functions on open Riemann surfaces (\rightarrow 367 Riemann Surfaces). The value distribution of a set of several meromorphic functions was studied first by A. Bloch and developed into the study of **meromorphic curves** by Ahlfors, and H. and J. Weyl [2]. The behavior of meromorphic functions in neighborhoods of general singularities has also been studied. An example of results in that field is the theory of \dagger cluster sets.

References

- [1] R. H. Nevanlinna, *Eindeutige analytische Funktionen*, Springer, 1935, revised edition, 1953; English translation, *Analytic functions*, Springer, 1970.

[2] H. Weyl and J. Weyl, *Meromorphic functions and analytic curves*, Princeton Univ. Press, 1943.

[3] G. Valiron, *Directions de Borel des fonctions méromorphes*, *Mémor. Sci. Math.*, Gauthier-Villars, 1938.

[4] W. K. Hayman, *Meromorphic functions*, Clarendon Press, 1964.

[5] W. Fuchs, *Developments in the classical Nevanlinna theory of meromorphic functions*, *Bull. Amer. Math. Soc.*, 73 (1967), 275–291.

[6] A. Edrei and W. Fuchs, *The deficiencies of meromorphic functions of order less than one*, *Duke Math. J.*, 27 (1960), 233–249.

[7] S. Hellerstein and J. Williamson, *Entire functions with negative zeros and a problem of R. Nevanlinna*, *J. Anal. Math.*, 22 (1969), 233–267.

[8] A. Edrei, *Solution of the deficiency problem for functions of small order*, *Proc. London Math. Soc.*, 26 (1973), 435–445.

[9] A. Weitsman, *Meromorphic functions with maximal deficiency sum and a conjecture of F. Nevanlinna*, *Acta Math.*, 123 (1969), 115–139.

[10] A. Weitsman, *A theorem on Nevanlinna deficiencies*, *Acta Math.*, 128 (1972), 41–52.

[11] D. Drasin, *The inverse problem of Nevanlinna theory*, *Acta Math.*, 138 (1977), 83–151.

[12] A. Gol'dberg, *Meromorphic functions*, *J. Soviet Math.*, 4 (1975), 157–216.

[13] V. Petrenko, *Growth of meromorphic functions of finite lower order (in Russian)*, *Izv. Akad. Nauk SSSR*, 33 (1969), 414–454.

Also \rightarrow references to 124 *Distribution of Values of Functions of a Complex Variable*.

273 (II.17) Metric Spaces

A. General Remarks

The distance between two points $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in the n -dimensional \dagger Euclidean space \mathbf{R}^n is defined by $\rho(x, y) = \sqrt{(y_1 - x_1)^2 + \dots + (y_n - x_n)^2}$. The function $\rho(x, y)$ is nonnegative for every pair (x, y) and has the following properties: (i) $\rho(x, y) = 0$ if and only if $x = y$; (ii) $\rho(x, y) = \rho(y, x)$; and (iii) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ for any three points x, y, z . Property (iii) is called the **triangle inequality**.

B. Definition of Metric Spaces

Abstracting the notion of distance from Euclidean spaces, M. Fréchet defined metric

spaces [1] (1906). A **metric** on a set X is a nonnegative function ρ on $X \times X$ that satisfies (i), (ii), and (iii) of Section A, and a **metric space** (X, ρ) , or simply X , is a set X provided with a metric ρ . The members of X are called points, ρ is called the **distance function**, and $\rho(x, y)$ is called the **distance** from x to y . The distance function is sometimes denoted by $d(x, y)$ or $\text{dis}(x, y)$. If (i) is replaced by its weaker form (i') $\rho(x, x) = 0$, the function ρ is called a **pseudometric** (or **pseudodistance function**), and X is called a **pseudometric space**.

Examples of metric spaces:

(1) The n -dimensional Euclidean space \mathbf{R}^n , in particular the real number system \mathbf{R} with $\rho_0(x, y) = |x - y|$. (2) The \dagger function space $L_p(\Omega)$. (3) The \dagger function space $C(\Omega)$. (4) The \dagger sequence space s , i.e., the space $\mathbf{R}^{\mathbb{N}}$ of all sequences of real numbers with metric $\rho(x, y) = \sum_{n=1}^{\infty} 2^{-n} |x_n - y_n| / (1 + |x_n - y_n|)$, where $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$. (5) The \dagger sequence space m , i.e., the space of all bounded sequences of real numbers with metric $\rho(x, y) = \sup_n |x_n - y_n|$ for $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$. (6) A **Baire zero-dimensional space** $(\Omega^{\mathbb{N}}, \rho)$, where Ω is a set and the distance $\rho(x, y)$ between $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ is equal to the reciprocal of the minimum n such that $x_n \neq y_n$. When the cardinal number τ of Ω is specified, the Baire space is denoted by $B(\tau)$. (7) For any set X , define ρ by setting $\rho(x, x) = 0$ and $\rho(x, y) = 1$ when $x \neq y$. Then (X, ρ) is a metric space, called a **discrete metric space**. (8) For any set X , define ρ by setting $\rho(x, y) = 0$ for any members x and y . Then ρ is a pseudometric, and the resulting space X is called an **indiscrete pseudometric space**.

For a subset M of a metric space X , $\sup\{\rho(x, y) \mid x, y \in M\}$ is called the **diameter** of M (denoted by $d(M)$), and M is said to be **bounded** if its diameter is finite (including $M = \emptyset$). For two subsets A, B of X , $\inf\{\rho(x, y) \mid x \in A, y \in B\}$ is called the **distance** between A and B , denoted by $\rho(A, B)$. We have $\rho(A, B) = \rho(B, A)$. When a family $\mathfrak{M} = \{M_\lambda \mid \lambda \in \Lambda\}$ of subsets of X is a covering of X , i.e., $X = \bigcup_\lambda M_\lambda$, the supremum of the diameters $d(M_\lambda)$ of M_λ in \mathfrak{M} , $\sup\{d(M_\lambda) \mid \lambda \in \Lambda\}$, called the **mesh of the covering** \mathfrak{M} . For a positive number ε , a covering whose mesh is less than ε is called an ε -**covering**. A metric space X is called **totally bounded** (or **precompact**) (F. Hausdorff, 1927) if for each positive number ε there exists a finite ε -covering of X .

A subset X_1 of a metric space X becomes a metric space if we define its metric ρ_1 by setting $\rho_1(x, y) = \rho(x, y)$ for $x, y \in X_1$, where ρ is the metric of X . The space (X_1, ρ_1) is called a **metric subspace** of (X, ρ) . A subset of X is called **totally bounded** (or **precompact**) if it is

totally bounded as a metric subspace. Any totally bounded subset is bounded. Conversely, in the Euclidean space \mathbf{R}^n any bounded subset is totally bounded.

A bijection f from a metric space (X_1, ρ_1) onto a metric space (X_2, ρ_2) is called an **isometric mapping** if f preserves the metric, i.e., $\rho_2(f(x), f(y)) = \rho_1(x, y)$ for any points $x, y \in X_1$; and X_1 and X_2 are called **isometric** if there is an isometric mapping from X_1 onto X_2 .

Let X be a metric space with metric ρ , and let f be an injection from a set Y into X . Then the function $\rho'(y_1, y_2) = \rho(f(y_1), f(y_2))$ ($y_1, y_2 \in Y$) is a distance function on Y , and with this metric the set Y becomes a metric space called the **metric space induced by f** ; f is an isometric mapping from (Y, ρ') onto $(f(Y), \rho)$.

For a finite number of metric spaces $(X_1, \rho_1), \dots, (X_n, \rho_n)$, we can define a metric ρ on their Cartesian product $X = X_1 \times \dots \times X_n$ by setting

$$\rho(x, y) = \sqrt{\rho_1(x_1, y_1)^2 + \dots + \rho_n(x_n, y_n)^2},$$

for two points $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ of X . Thus we obtain a metric space (X, ρ) , called the **product metric space** of $(X_1, \rho_1), \dots, (X_n, \rho_n)$. The n -dimensional Euclidean space \mathbf{R}^n is the product metric space of n copies of the real line (\mathbf{R}, ρ_0) .

C. Topology for Metric Spaces

For a point x of a metric space (X, ρ) and any positive number ε , the set $U_\varepsilon(x)$ of all points y such that $\rho(x, y) < \varepsilon$ is called the ε -**neighborhood** (or ε -**sphere**) of x . We can introduce a topology for X by taking the family of all ε -neighborhoods as a \dagger base for the neighborhood system (\rightarrow 425 Topological Spaces). Then the following five propositions hold, any one of which can be used to define the same topology: (i) A subset O is \dagger open if and only if for any point x in O there is a positive number ε such that the ε -neighborhood of x is contained in O . (ii) A subset F is \dagger closed if and only if any point whose every ε -neighborhood contains at least one point of F is contained in F . (iii) A subset U is a neighborhood of a point x if and only if U contains some ε -neighborhood of x . (iv) A point x is an \dagger interior point of a subset A if and only if some ε -neighborhood of x is contained in A ; the interior A^i of A is the set of all such points. (v) A point x is adherent to a subset A if every ε -neighborhood of x contains at least one point of A ; the closure \bar{A} of A is the set of all such points, and $x \in \bar{A}$ if and only if $\rho(x, A) = 0$.

Every metric space X satisfies the \dagger first countability axiom. A metric space X is a

*Hausdorff space and, more specifically, a *perfectly normal space; it is also *paracompact.

In the same way, we can define a topology for each pseudometric space that satisfies the first countability axiom, but a pseudometric space is not necessarily Hausdorff.

D. Convergence of Sequences

A sequence $\{x_n\}$ of points in a metric space is said to **converge** to a point x (written $\lim_{n \rightarrow \infty} x_n = x$) if $\rho(x_n, x)$ tends to zero as $n \rightarrow \infty$. The point x is called the **limit** of $\{x_n\}$. This convergence is equivalent to convergence with respect to the topology defined in Section C (\rightarrow 87 Convergence). As the first countability axiom is satisfied, we may define the topology by means of convergent sequences of points: the closure \bar{A} of a subset A is the set of all limits of sequences of points in A .

E. Separable Metric Spaces

For a metric space X the following three conditions are equivalent: (i) There exists a countable family \mathfrak{D}_0 of open sets of X such that each open set of X is the union of members of \mathfrak{D}_0 (*second countability axiom). (ii) X is *separable, that is, X has a countable subset that is *dense in X . (iii) Every open covering of X has a countable subcovering (*Lindelöf space). A metric space with any of these properties is called a **separable metric space**. The sequence space s is separable. Any separable metric space can be isometrically embedded in the sequence space m , i.e., is isometric to a subspace of m (\rightarrow 168 Function Spaces B).

F. Compact Metric Spaces

For a metric space X , the following five conditions are equivalent: (i) X is *compact, that is, every open covering of X has a finite subcovering. (ii) X is *countably compact, that is, every countable open covering of X has a finite subcovering. (iii) X is *sequentially compact, that is, any sequence of points in X has a convergent subsequence. (iv) Every nested family $F_1 \supset F_2 \supset \dots$ of nonempty closed sets of X has a nonempty intersection. (v) Every infinite subset M of X has an accumulation point x , i.e., $x \in \overline{M - \{x\}}$. A metric space satisfying any of these conditions is called a **compact metric space** (M. Fréchet [1]). Every real-valued continuous function defined on a compact metric space has a maximum and a minimum. A metric space is compact if and only if it is totally bounded and complete (\rightarrow Section J). Every totally bounded metric space is

separable. In particular, every compact metric space is separable.

Let $\mathfrak{U} = \{U_\lambda\}$ be an open covering of a compact metric space X . There exists a positive number δ such that every set with $d(A) < \delta$ is contained in some U_λ . The number δ is called the **Lebesgue number** of the open covering \mathfrak{U} .

A subset A of a metric space is said to be **compact** if it is compact as a metric subspace, and A is said to be **relatively compact** if its closure is compact. Bounded closed sets in \mathbf{R}^n , in particular closed intervals of real numbers, are compact. For these sets, conditions (i), (iv), and (v) are called the **Heine-Borel theorem** (or **Borel-Lebesgue theorem**), **Cantor's intersection theorem**, and the **Bolzano-Weierstrass theorem**, respectively.

G. Product Spaces of Metric Spaces

Let $(X_1, \rho_1), \dots, (X_n, \rho_n)$ be metric spaces. Then the Cartesian product $X = X_1 \times \dots \times X_n$ has distance functions

$$\bar{\rho}_p(x, y) = \{\rho_1(x_1, y_1)^p + \dots + \rho_n(x_n, y_n)^p\}^{1/p},$$

$p \geq 1$

and

$$\bar{\rho}_\infty(x, y) = \max\{\rho_1(x_1, y_1), \dots, \rho_n(x_n, y_n)\},$$

where $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. The topology of X induced by each one of these metrics coincides with the product topology. In particular, for the n -dimensional Euclidean space \mathbf{R}^n , the metrics $\bar{\rho}_p$ ($p \geq 1$) and $\bar{\rho}_\infty$ define the same topology.

Let $(X_1, \rho_1), \dots, (X_n, \rho_n), \dots$ be a countable number of metric spaces. If we define a metric ρ on the Cartesian product $X = \prod_{n=1}^\infty X_n$ by

$$\rho(x, y) = \sum_{n=1}^\infty \frac{1}{2^n} \frac{\rho_n(x_n, y_n)}{1 + \rho_n(x_n, y_n)},$$

where $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$, then the topology defined by ρ is identical with the product topology. For the Cartesian product of an uncountable number of metric spaces, we cannot construct a metric ρ such that the topology induced by ρ agrees with the product topology in general.

H. Uniformity of Metric Spaces

Every metric space X is a *uniform space, for which we may take a countable number of subsets $\{(x, y) | \rho(x, y) < 2^{-n}\}$, $n = 1, 2, \dots$, of $X \times X$ as a base of *uniformity (\rightarrow 436 Uniform Spaces).

I. Uniform Continuity

A mapping f from a metric space (X, ρ) into a metric space (Y, σ) is \dagger continuous if for any point x in X and any positive number ε there is a positive number δ such that $f(U_\delta(x)) \subset V_\varepsilon(f(x))$, where $U_\delta(x)$ is a δ -neighborhood for ρ and $V_\varepsilon(y)$ is an ε -neighborhood for σ ; that is, $\rho(x, x') < \delta$ implies $\sigma(f(x), f(x')) < \varepsilon$. In this case, we must generally choose δ depending on x and ε . In the special case where we can choose δ depending only on ε , independently of x , we call f **uniformly continuous** in X . (The notion of uniform continuity may be generalized to uniform spaces.) Not every continuous mapping is necessarily uniformly continuous, but every continuous mapping from a compact metric space into a metric space is uniformly continuous.

J. Complete Metric Spaces

A sequence $\{x_n\}$ of points in a metric space (X, ρ) is called a **fundamental sequence** (or **Cauchy sequence**) if $\rho(x_n, x_m) \rightarrow 0$ as $n, m \rightarrow \infty$. Every convergent sequence is a fundamental sequence, but the converse is not always true. A metric space is called **complete** if every fundamental sequence in the space converges to some point of the space (M. Fréchet [1]). A topological space that is homeomorphic with a complete separable metric space is sometimes called a **Polish space** (\rightarrow 22 Analytic Sets I). The metric spaces introduced in examples (1) through (5) of Section B are complete. (In example (3) we must assume that the space Ω is a compact Hausdorff space.) A metric space is compact if and only if it is complete and totally bounded. A \dagger locally compact metric space is homeomorphic to a complete metric space.

For a metric space X , we can construct a complete metric space Y such that there is an isometric mapping φ from X onto a dense subspace X_1 of Y (F. Hausdorff, 1914). Such a pair (Y, φ) is called a **completion** of X . If X has two completions (Y_1, φ_1) and (Y_2, φ_2) , then there is an isometric mapping f from Y_1 onto Y_2 with $\varphi_2 = f \circ \varphi_1$. In this sense the completion of X is unique. By identifying X with $\varphi(X)$ when (Y, φ) is the completion of X , any metric space can be regarded as a dense subspace of a complete metric space. For example, the completion of the rational number system \mathbf{Q} is the real number system \mathbf{R} . A metric space is totally bounded ($=$ precompact) if and only if its completion is compact.

Baire-Hausdorff theorem: In a complete metric space every set of the \dagger first category is a \dagger boundary set. That is, every set that can be

expressed as the union of a countable number of sets whose closures have no interior point has no interior point. In other words, if the union $\bigcup_{n=1}^{\infty} F_n$ of closed sets F_1, F_2, \dots of X has an interior point, then at least one of the F_n must have an interior point.

K. The Metrization Problem

A topological space X is called **metrizable** if we can introduce a suitable metric for X which induces a topology identical to the original one. A $\dagger T_1$ -space satisfying the second countability axiom is metrizable if and only if it is \dagger regular (**Uryson-Tikhonov theorem**; P. S. Uryson, *Math. Ann.*, 94 (1925), A. Tikhonov, *Math. Ann.*, 95 (1925)). However, a metric space does not necessarily satisfy the second countability axiom. Therefore, the Uryson-Tikhonov theorem does not provide a necessary and sufficient condition for metrizability. The following are some necessary and sufficient conditions for a topological space X to be metrizable:

(1) There exists a nonnegative real-valued function d on $X \times X$ satisfying the first two axioms given in Section A and the following condition: There exists a real-valued function $\varphi(w)$ that converges to zero as $w \rightarrow 0$ such that, for any three points x, y, z and any positive number ε , $d(x, y) < \varphi(\varepsilon)$ and $d(y, z) < \varphi(\varepsilon)$ imply $d(x, z) < \varepsilon$ (E. W. Chittenden, *Trans. Amer. Math. Soc.*, 18 (1917)).

(2) X is a T_1 -space that has a countable number of open coverings $\mathfrak{M}_1, \mathfrak{M}_2, \dots$ satisfying the following two conditions: (i) If $U_1, U_2 \in \mathfrak{M}_{n+1}$ have a common point, there is a set $U \in \mathfrak{M}_n$ with $U \supset U_1 \cup U_2$; (ii) for any point x in X , if U_n is any member of \mathfrak{M}_n containing x , the family $\{U_n\}_{n=1,2,\dots}$ is a base for the neighborhood system of x (P. S. Aleksandrov and Uryson, *C. R. Acad. Sci., Paris*, 177 (1923), and N. Aronszajn). When X is a uniform space, this amounts to saying that X has a metric compatible with the uniform structure if and only if X is a T_1 -space and has a countable base of uniformity.

(3) X is a T_1 -space that admits a countable number of open coverings $\mathfrak{M}_1, \mathfrak{M}_2, \dots$ such that $\{S(S(x, \mathfrak{M}_i), \mathfrak{M}_j) | i, j = 1, 2, \dots\}$ is a base for the neighborhood system of x at each point of X , where $S(A, \mathfrak{M})$ is the \dagger star of A relative to \mathfrak{M} (R. L. Moore, *Fund. Math.*, 25 (1935); K. Morita, *Proc. Japan Acad.*, 27 (1951); A. H. Stone, *Pacific J. Math.*, 10 (1960); A. V. Arkhangel'skii, *Dokl. Akad. Nauk SSSR*, 2 (1961)).

(4) X is regular and has a $\dagger \sigma$ -locally finite open base (J. Nagata, *J. Inst. Polytech. Osaka City Univ.*, 1 (1950); Yu. M. Smirnov, *Uspekhi Mat. Nauk*, 6 (1951)).

(5) X is regular and has a $\dagger\sigma$ -discrete open base (R. H. Bing, *Canad. J. Math.*, 3 (1951)).

(6) X is a \dagger collectionwise normal Moore space (\rightarrow below; Bing, *ibid.*).

(7) X is a \dagger perfect image of a subspace of a Baire's zero-dimensional space (Morita, *Sci. Rep. Tokyo Kyoiku Daigaku*, sec. A, 5 (1955)).

(8) X is a Hausdorff $\dagger M$ -space such that the diagonal is a $\dagger G_\delta$ -set in the direct product $X \times X$ (A. Okuyama, *Proc. Japan Acad.* 40 (1964); C. J. R. Borges, *Pacific J. Math.*, 17 (1966); J. Chaber, *Fund. Math.*, 94 (1977)).

A regular space is said to be a Moore space if it has a countable number of open coverings \mathfrak{M}_i such that $\{S(x, \mathfrak{M}_i)\}$ is a base for the neighborhood system of x for any point x . A Moore space is not necessarily metrizable. F. B. Jones (*Bull. Amer. Math. Soc.*, 43 (1937)) proved under the assumption $2^{\aleph_0} < 2^{\aleph_1}$ that every separable normal Moore space is metrizable and asked whether or not every normal Moore space is metrizable. (3) and (6) are partial answers to the question. A normal Moore space is metrizable if it is locally compact and \dagger locally connected (G. Reed and P. Zenor). The existence of a nonmetrizable separable normal Moore space is consistent with and independent of the axioms of the usual ZFC set theory, the \dagger Zermelo-Fraenkel set theory with the \dagger axiom of choice (F. D. Tall). W. G. Fleissner (*Trans. Amer. Math. Soc.*, 273 (1982)) constructed a normal nonmetrizable Moore space assuming an axiom weaker than the continuum hypothesis, while P. J. Nyikos (1980) has proved that every normal space with the first countability axiom is collectionwise normal from the strong axiom of set theory.

In connection with (7) the following result is known. Let f be a \dagger closed continuous mapping from a metric space X onto a topological space Y . Then the following conditions are equivalent: (1) Y is metrizable; (2) For each $y \in Y$ the \dagger boundary $\partial f^{-1}(y)$ of the inverse image is compact; (3) Y satisfies the first countability axiom (Morita and S. Hanai, *Proc. Japan Acad.*, 32 (1956); Stone, *Proc. Amer. Math. Soc.*, 7 (1956); I. A. Vainstein, *Dokl. Akad. Nauk SSSR*, 57 (1947)). In particular, perfect images of metric spaces are metrizable.

For quotient topological spaces of metric spaces \rightarrow 425 Topological Spaces CC.

References

[1] M. Fréchet, Sur quelques points du calcul fonctionnel, *Rend. Circ. Mat. Palermo*, 22 (1906), 1–74.

Also \rightarrow references to 425 Topological Spaces. For metrization problem see in particular

[2] J. Nagata, *Modern general topology*, North-Holland, 1968.

[3] R. Engelking, *General topology*, Polish Scientific Publishers, 1977.

[4] Y. Kodama and K. Nagami, *Theory of topological spaces (in Japanese)*, Iwanami, 1974.

[5] A. V. Arkhangel'skiĭ, Mappings and spaces, *Russian Math. Surveys*, 21 (1966), 115–162.

274 (XII.16) Microlocal Analysis

A. General Remarks

Let X be an open set of \mathbf{R}^n . Then $X \times (\mathbf{R}^n \setminus 0)$ can be identified with $T^*X \setminus 0$, the \dagger cotangent bundle of X minus the zero-section. To every locally integrable function (or distribution or hyperfunction (\rightarrow 125 Distributions and Hyperfunctions)) $f(x)$ defined on X , one can assign closed subsets of $X \times (\mathbf{R}^n \setminus 0)$ called the wave front set of f and the singularity spectrum (or analytic wave front set or essential support) of f . The wave front set (resp. the singularity spectrum) of f describes in detail the singularity of f modulo the infinitely differentiable functions (resp. the real analytic functions). One can further associate with f more refined objects defined on $X \times (\mathbf{R}^n \setminus 0)$, such as microfunctions. In many cases one can recover knowledge of the structure of f by analyzing these objects defined on $X \times (\mathbf{R}^n \setminus 0)$. Such an analysis on the cotangent bundle of X is called **microlocal analysis**. Microlocal analysis is particularly successful if f is a solution of a system of linear (pseudo-)differential equations, because in that case one can use various linear transformations, such as differential operators, pseudodifferential operators (\rightarrow 345 Pseudodifferential Operators), or microdifferential operators and Fourier integral operators, or quantized contact transformations.

B. Microlocal Analysis for Distributions

Let u be a distribution defined in an open subset X of \mathbf{R}^n . The **wave front set** $\text{WF}(u)$ of u is defined as the complement in $X \times (\mathbf{R}^n \setminus 0)$ of the collection of all (x_0, ξ_0) in $X \times (\mathbf{R}^n \setminus 0)$ such that for some neighborhood U of x_0 , V of ξ_0 we have for each $\varphi \in C_0^\infty(U)$ and each $N > 0$,

$$\langle u, \varphi \exp(-i\tau x \cdot \xi) \rangle = O(\tau^{-N})$$

as $\tau \rightarrow \infty$, uniformly in $\xi \in V$ (L. Hörmander; [1]). $\text{WF}(u)$ is considered to be a subset of

$T^*X \setminus \{0\}$. If $\text{WF}(u) = \emptyset$, then u is a C^∞ -function. Let $\pi: X \times (\mathbf{R}^n \setminus 0) \rightarrow X$ be the natural projection. If $\pi(\text{WF}(u))$ contains x_0 , then u is not a C^∞ -function in any small neighborhood of x_0 . Thus $\text{WF}(u)$ is the obstruction for u to be infinitely differentiable. Let $p(x, D)$ be a \dagger linear partial differential operator of order m . Assume that $p(x, D)u = f$. Then

$$\text{WF}(f) \subset \text{WF}(u) \subset \text{WF}(f) \cup p_m^{-1}(0),$$

where $p_m: (x, \xi) \rightarrow p_m(x, \xi)$ is the \dagger principal symbol of $p(x, D)$. Such a technique of localizing the problem on the cotangent bundle has been used in the form of the estimation of the Fourier transform of f since the advent of the \dagger singular integral operators of A. P. Calderón and A. Zygmund [3, 4] (see also S. Mizohata [5, 6]). The formula above contains as a special case the classical result that u is a C^∞ -function if $p(x, D)$ is \dagger elliptic and if f is a C^∞ -function.

In obtaining useful results of microlocal analysis for distributions one often uses Fourier integral operators and pseudodifferential operators (or singular integral operators) (→ 345 Pseudodifferential Operators).

C. Fourier Integral Operators [1, 2, 7–9]

A Fourier integral operator $B: C_0^\infty(\mathbf{R}^n) \rightarrow \mathcal{D}'(\mathbf{R}^n)$ is a locally finite sum of linear operators of the type

$$Af(x) = (2\pi)^{-(n+N)/2} \int_{\mathbf{R}^{n+N}} a(x, \theta, y) \times \exp(i\varphi(x, \theta, y)) f(y) dy d\theta.$$

Here $a(x, \theta, y)$ is a C^∞ -function satisfying the inequality

$$|D_x^\alpha D_\theta^\beta D_y^\gamma a(x, \theta, y)| \leq C(1 + |\theta|)^{m-\rho|\beta| + (1-\rho)(|\alpha| + |\gamma|)}$$

for some fixed m and ρ , $1 \geq \rho > 1/2$, and any triple of \dagger multi-indices α, β, γ , and $\varphi(x, \theta, y)$ is a real-valued C^∞ -function which is homogeneous of degree 1 in θ for $|\theta| > 1$. The function φ is called the **phase function** and a the **amplitude function**.

Let $C_\varphi = \{(x, \theta, y) | d_\theta \varphi(x, \theta, y) = 0, \theta \neq 0\}$ and $W = \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^n | \exists \theta \neq 0 \text{ such that } (x, \theta, y) \in C_\varphi\}$. If $d_{x, \theta, y} \varphi(x, \theta, y) \neq 0$ for $\theta \neq 0$, then the kernel distribution $k(x, y)$ of A is of class C^∞ outside W . A phase function φ is called nondegenerate if the $d_{x, \theta, y}(\partial \varphi(x, \theta, y) / \partial \theta_j)$, $j = 1, 2, \dots, N$, are linearly independent at every point of C_φ . In this case, C_φ is a smooth manifold in \mathbf{R}^{n+N+n} and the mapping $\Phi: C_\varphi \ni (x, \theta, y) \rightarrow (x, y, \xi, \eta)$, $\xi = d_x \varphi(x, \theta, y)$, $\eta = d_y \varphi(x, \theta, y)$, is an immersion of C_φ to $T^*(\mathbf{R}^n \times \mathbf{R}^n) \setminus 0$, the cotangent bundle of $\mathbf{R}^n \times \mathbf{R}^n$ minus its zero-section. The image $\Phi C_\varphi = \Lambda_\varphi$ is a **conic Lagrangian manifold**, i.e., the canonical 2-form σ

$= \sum_j d\xi_j \wedge dx_j - \sum_j d\eta_j \wedge dy_j$ vanishes on Λ_φ and the multiplicative group of positive numbers acts on Λ_φ . Let $\lambda_1, \lambda_2, \dots, \lambda_{2n}$ be a system of local coordinates in Λ_φ . These, together with $\partial \varphi / \partial \theta_1, \partial \varphi / \partial \theta_2, \dots, \partial \varphi / \partial \theta_N$, constitute a system of local coordinate functions of \mathbf{R}^{n+N+n} in a neighborhood of C_φ . Let J denote the Jacobian determinant

$$\frac{D\left(\lambda_1, \lambda_2, \dots, \lambda_{2n}, \frac{\partial \varphi}{\partial \theta_1}, \dots, \frac{\partial \varphi}{\partial \theta_N}\right)}{D(x, \theta, y)}.$$

The function $a_{\Lambda_\varphi} = \sqrt{J} a|_{C_\varphi} \Phi^{-1} \exp(\pi Mi/4)$ is called the **symbol** of A . Here $a|_{C_\varphi}$ is the restriction of a to C_φ and M is an integer called the **Keller-Maslov index** [10, 11]. The conic Lagrangian manifold $\Lambda_\varphi = \Lambda_\varphi(A)$ and the symbol $a_{\Lambda_\varphi} = a_{\Lambda_\varphi}(A)$ essentially determine the singularity of the \dagger kernel distribution $k(x, y)$ of the Fourier integral operator A . Conversely, given a conic Lagrangian manifold Λ in $T^*(\mathbf{R}^n \times \mathbf{R}^n) \setminus 0$ and a function a_Λ on it, one can construct locally a Fourier integral operator A such that $\Lambda_\varphi(A) = \Lambda$ and $a_{\Lambda_\varphi}(A) = a_\Lambda$. For global construction of such a Fourier integral operator one requires detailed consideration of the Keller-Maslov index. A globally defined Fourier integral operator A with $\Lambda_\varphi(A) = \Lambda$ and $a_{\Lambda_\varphi}(A) = a_\Lambda$ exists if and only if a_Λ is not a function on Λ but a section of the complex line bundle $\Omega_{1/2} \otimes L$, where $\Omega_{1/2}$ is the bundle of square roots of the volume elements of Λ and L is a \mathbf{Z}_4 bundle over Λ called the **Maslov bundle**. The factor $\sqrt{J} \exp(\pi Mi/4)$ in the definition of a_{Λ_φ} above appears as the effect of trivialization of the bundle $\Omega_{1/2} \otimes L$. Those Fourier integral operators whose associated conic Lagrangian manifolds are the graphs of \dagger homogeneous canonical transformations of $T^*(\mathbf{R}^n)$ are most frequently used in the theory of linear partial differential equations. Let A be a Fourier integral operator such that $\Lambda_\varphi(A)$ is the graph of a homogeneous canonical transformation χ . Then the adjoint of A is a Fourier integral operator such that the associated conic Lagrangian manifold is the graph of the inverse transformation χ^{-1} . Let A_1 be another such operator; if $\Lambda_\varphi(A_1)$ is the graph of χ_1 , then the composed operator $A_1 A$ is also a Fourier integral operator and $\Lambda_\varphi(A_1 A)$ is the graph of the composed homogeneous canonical transformation $\chi_1 \chi$.

Consider the kernel distribution $k(x, y)$ of A . If the phase function φ of A is nondegenerate, then $\text{WF}(k)$ is contained in $\Lambda_\varphi(A)$. Moreover, if the symbol $a_{\Lambda_\varphi}(A)$ does not vanish, then $\text{WF}(k) = \Lambda_\varphi(A)$. Let u be a distribution and A be a Fourier integral operator such that $\Lambda_\varphi(A)$ is the graph of a homogeneous canonical transformation χ . Then $\text{WF}(Au) \subset \chi(\text{WF}(u))$.

A pseudodifferential operator of class $S_{\rho,1-\rho}^m(\mathbf{R}^n)$ is a particular type of Fourier integral operator (\rightarrow 345 Pseudodifferential Operators). In fact, a Fourier integral operator A is a pseudodifferential operator of class $S_{\rho,1-\rho}^m(\mathbf{R}^n)$ if and only if $\Lambda_\varphi(A)$ is the graph of the identity mapping of $T^*(\mathbf{R}^n)$. Hence for any Fourier integral operator A , A^*A and AA^* are pseudodifferential operators.

The following theorem is due to Yu. V. Egorov [12]: Let $P(x, D)$ be a pseudodifferential operator of class $S_{\rho,1-\rho}^m(\mathbf{R}^n)$ with the symbol $p(x, \xi)$, and let A be a Fourier integral operator such that the associated conic Lagrangian manifold $\Lambda_\varphi(A)$ is the graph of a homogeneous canonical transformation χ of $T^*(\mathbf{R}^n)$. Then there exists a pseudodifferential operator $Q(x, D)$ with the symbol $q(x, \xi) \in S_{\rho,1-\rho}^m(\mathbf{R}^n)$ such that $P(x, D)A = AQ(x, D)$ and $q(x, \xi) - p(\chi(x, \xi)) \in S_{\rho,1-\rho}^{m-2\rho+1}(\mathbf{R}^n)$. Note that $m - 2\rho + 1 < m$.

Assume that $m = 1$, $\rho = 1$, and that $p_1(x, \xi)$ is a real-valued C^∞ -function, homogeneous of degree 1 in ξ for $|\xi| > 1$, such that $p(x, \xi) - p_1(x, \xi) \in S_{1,0}^0(\mathbf{R}^n)$ and $d_\xi p_1(x^0, \xi^0) \neq 0$ at (x^0, ξ^0) , where $p_1(x^0, \xi^0) = 0$. Then one can find a Fourier integral operator A such that the function $q(x, \xi)$ of Egorov's theorem satisfies the relation $q(x, \xi) - \xi_1 \in S_{1,0}^0(\mathbf{R}^n)$.

The boundedness of Fourier integral operators in the space $L_2(\mathbf{R}^n)$ (or the spaces $H^s(\mathbf{R}^n)$) has also been studied in several cases. Some sufficient conditions for boundedness can be found in [7, 8, 14–16].

The theory of Fourier integral operators has its origin in the asymptotic representation of solutions of the wave equation, (see, e.g., [17, 18]). Fourier integral operators were first used by G. I. Eskin [7].

D. Essential Support or Analytic Wave Front Set of a Distribution

Inspired by the physical idea introduced by C. Chandler and H. P. Stapp, J. Bros and D. Iagolnitzer introduced the notion of the **essential support** of a distribution, which is a closed subset of $X \times (\mathbf{R}^n \setminus 0)$ [19]. Let u be a distribution defined on an open set X of \mathbf{R}^n and χ be a C^∞ -function with compact support around $x_0 \in X$ which is locally analytic and different from 0 at x_0 . Let $\Sigma_\chi(u)$ be the subset of $\mathbf{R}^n \setminus 0$ of which the complement is defined as follows. A point η is in the complement of $\Sigma_\chi(u)$ if there exist a conic neighborhood U of η , constants $\alpha, \gamma_0 > 0$, and C_N such that

$$|\langle u, \chi \exp\{-ix'\xi - \gamma|\xi||x - x_0|^2\} \rangle| \leq C_N(1 + |\xi|)^{-N} e^{-\alpha\gamma|\xi|}$$

for all $\xi \in U$, $0 < \gamma < \gamma_0$, and all positive integers N .

The essential support $\Sigma_{x_0}(u)$ of u at x_0 , is the limit of $\Sigma_\chi(u)$ when the width $|\text{supp } \chi|$ of the support of χ around x_0 tends to 0. The **essential support** $\Sigma(u)$ of u is the closed subset $\bigcup_{x \in X} \{x\} \times \Sigma_x(u)$ of $X \times \mathbf{R}^n \setminus 0$. $\Sigma(u)$ is the obstruction for u to be real analytic. L. Hörmander [20] also defined the **analytic wave front set** of a distribution, which is also the obstruction for a distribution to be real analytic. The definition of an analytic wave front set is quite different from that of essential support. However they coincide with each other [21]. Moreover, both of them coincide with the singularity spectrum of u if the distribution u is regarded as a hyperfunction.

E. Microlocal Analysis for Hyperfunctions [22]

(1) Microfunctions. Let N be a real analytic manifold of dimension $n + d$ and M its submanifold of codimension d . In what follows, $T_M N$ and $T_M^* N$ denote the normal bundle of M and the conormal bundle supported by M , respectively. Here the **normal bundle** $T_M N$ is defined to be the quotient bundle $TN|_M / TM$ of the tangent bundle and the **conormal bundle** $T_M^* N$ to be the subbundle of the cotangent bundle $T^* N|_M$ that annihilates TM . Identifying N with $\{(x, v) \in TN | v = 0\}$ or $\{(x, \xi) \in T^* N | \xi = 0\}$, we define the **tangential sphere bundle** SN and **cotangential sphere bundle** $S^* N$ by $(TN \setminus N)/R_+^\times (= \bigcup_{x \in N} (T_x N \setminus \{0\})/R_+^\times)$ and $(T^* N \setminus N)/R_+^\times (= \bigcup_{x \in N} (T_x^* N \setminus \{0\})/R_+^\times)$, respectively. The **normal sphere bundle** $S_M N = (T_M N \setminus M)/R_+^\times$ and **conormal sphere bundle** $S_M^* N = (T_M^* N \setminus M)/R_+^\times$ are defined in the same manner. In parallel with the algebraic geometry (\rightarrow 16 Algebraic Varieties) we define the **real monoidal transform** of N with center M to be the manifold $(N \setminus M) \cup S_M N$ with boundary, in which the center M is blown up to $S_M N$ by the polar coordinates. We denote it by $\tilde{M}N$. We mainly use this notion when N is a \ast -complexification X of M , regarding X as a $2n$ -dimensional real manifold. In this case we can canonically identify $T_M X$ with $\sqrt{-1} TM$, and hence $S_M X$ with $\sqrt{-1} SM$. We denote by $x + \sqrt{-1}v \neq 0$ the point in $S_M X$ that corresponds to $(x, \sqrt{-1}v)$ in $\sqrt{-1} SM$ by this identification. An open subset W of $X \setminus M$ is called a **conoidal neighborhood** of a subset U of $\sqrt{-1} SM$ if $W \cup \sqrt{-1} SM$ is a neighborhood of U in $\tilde{M}X$. Let $\varepsilon, \tilde{\varepsilon}$, and $\tilde{\tau}$ denote respectively the canonical embedding mappings from $X \setminus M$ to X , from $\tilde{M}X \setminus \sqrt{-1} SM$ to $\tilde{M}X$ and from $\sqrt{-1} SM$ to $\tilde{M}X$. We then define the \ast -sheaves $\tilde{\mathcal{O}}$ and $\tilde{\mathcal{O}}'$

by $\tilde{e}_* \varepsilon^{-1} \mathcal{O}_X$ and $\tilde{\mathcal{O}}|_{\sqrt{-1}SM} (= \tilde{\tau}^{-1} \tilde{\mathcal{O}}_X)$. Here \mathcal{O}_X denotes the sheaf of germs of holomorphic functions on X . We also define the sheaf \mathcal{Q} on $\sqrt{-1}SM$ by $\mathcal{H}^1|_{\sqrt{-1}SM}(\tau^{-1}\mathcal{O}_X)$, where τ is the canonical projection from $\tilde{M}X$ to X and \mathcal{H}_S^p denotes the p th \dagger derived functor of the functor Γ_S of taking the sections with support in S . Then \mathcal{Q} is isomorphic to $\tilde{\mathcal{A}}/\tau^{-1}\mathcal{A}$ for the sheaf \mathcal{A} of real analytic functions on M . The sheaf $\tilde{\mathcal{A}}$ is, so to speak, the sheaf of "boundary values" of holomorphic functions. Actually there exists a canonical mapping b from $\tilde{\mathcal{A}}$ to $\tau^{-1}\mathcal{B}_M$, where \mathcal{B}_M denotes the sheaf of \dagger hyperfunctions on M (\rightarrow 125 Distributions and Hyperfunctions). Thus we see that \mathcal{Q} describes the singularities of the boundary value of a holomorphic function. These sheaves $\tilde{\mathcal{A}}$ and \mathcal{Q} are easy to understand intuitively. However, they are defined on $\sqrt{-1}SM$, while $\sqrt{-1}S^*M$ is more important in analysis. Our final goal, namely, the sheaf of microfunctions, is constructed on $\sqrt{-1}S^*M$ through cohomological machinery starting from \mathcal{Q} . In order to do this, we introduce the disk bundle DM by

$$\{(x + \sqrt{-1}v0, (x, \sqrt{-1}\xi\infty))$$

$$\in \sqrt{-1}SM \times_M \sqrt{-1}S^*M | \langle v, \xi \rangle \leq 0\}.$$

Here $\sqrt{-1}SM \times_M \sqrt{-1}S^*M$ denotes the \dagger fiber product of $\sqrt{-1}SM$ and $\sqrt{-1}S^*M$ over M and the symbol $\xi\infty$ is used to emphasize that ξ designates the codirection, which is dual to the infinitesimally small quantity $v0$. The canonical projections from DM to $\sqrt{-1}S^*M$ and from $\sqrt{-1}SM$ to M are both denoted by τ . Similarly, the projections from DM to $\sqrt{-1}SM$ and from $\sqrt{-1}S^*M$ to M are denoted by π . We denote by a the antipodal mapping on $\sqrt{-1}S^*M$, namely, $a(x, \sqrt{-1}\xi\infty) = (x, -\sqrt{-1}\xi\infty)$. For a sheaf \mathcal{F} on $\sqrt{-1}S^*M$, we also denote $a_*\mathcal{F} (= a^{-1}\mathcal{F})$ by \mathcal{F}^a . In the following, $R^j\tau_*$, etc., denotes the j th \dagger right-derived functor of the functor τ_* of taking the direct image of sheaves, etc. (\rightarrow 383 Sheaves). Now the sheaf \mathcal{C}_M of **microfunctions** is defined on $\sqrt{-1}S^*M$ by $(R^{n-1}\tau_*\pi^{-1}\mathcal{Q})^a \otimes \pi^{-1}\omega_M$, where ω_M denotes the \dagger orientation sheaf of M . Note that $R^j\tau_*\pi^{-1}\mathcal{Q} = 0$ holds for $j \neq n-1$.

Remark: Here we have defined the sheaf \mathcal{C}_M of microfunctions on $\sqrt{-1}S^*M$. However, it is sometimes more convenient to define the sheaf on $\sqrt{-1}T^*M (= T^*M)$ by the following convention: $\hat{\mathcal{C}}_{M,(x,\sqrt{-1}\xi)} = \mathcal{C}_{M,(x,\sqrt{-1}\xi\infty)}$ if $\xi \neq 0$, and $\mathcal{B}_{M,x}$, if $\xi = 0$. Several authors (e.g., [23]) call this sheaf $\hat{\mathcal{C}}_M$ the sheaf of microfunctions and denote it by \mathcal{C}_M .

(2) Basic Properties of Microfunctions. The sheaf \mathcal{C}_M defined above has the following properties: (i) The sheaf \mathcal{C}_M is a \dagger flabby sheaf

on $\sqrt{-1}S^*M$. (ii) For each $(x, \sqrt{-1}\xi\infty)$ in $\sqrt{-1}S^*M$, there exists a surjective mapping from $\mathcal{B}_{M,x}$ to $\mathcal{C}_{M,(x,\sqrt{-1}\xi\infty)}$. This mapping is denoted by sp . The mapping from \mathcal{B}_M to $\pi_*\mathcal{C}_M$ is also denoted by sp . (iii) We have the following exact sequence: $0 \rightarrow \mathcal{A}_M \rightarrow \mathcal{B}_M \xrightarrow{sp} \pi_*\mathcal{C}_M \rightarrow 0$. (iv) $R^k\pi_*\mathcal{C}_M = 0$ holds for $k \neq 0$. The exact sequence (iii) shows that the singularities of hyperfunctions are dispersed over $\sqrt{-1}S^*M$ and that the dispersed object is described by the sheaf \mathcal{C}_M of microfunctions. For a hyperfunction $f \in \mathcal{B}_M(M)$, we call $sp(f) \in \mathcal{C}(\sqrt{-1}S^*M)$ the **spectrum** of f . We denote $\text{supp } sp(f)$ by $S.S.f$ and call it the **singularity spectrum** of f or the **singular spectrum**. It is known [21] that this coincides with the analytic wave front set of f and with the essential support of f if f is a distribution. (v) The following sequence is an exact sequence on $\sqrt{-1}SM$: $0 \rightarrow \tilde{\mathcal{A}}_M \xrightarrow{b} \tau^{-1}\mathcal{B}_M \rightarrow \pi_*\tau^{-1}\mathcal{C}_M \rightarrow 0$. In the following, a subset A of the $(n-1)$ -dimensional sphere S^{n-1} is said to be convex if $\varpi^{-1}(A) \cup \{0\}$ is convex, where ϖ is the canonical projection from $\mathbb{R}^n \setminus \{0\}$ to S^{n-1} , and if $\varpi^{-1}(A) \cup \{0\}$ is convex and includes no straight line, A is said to be **properly convex**. A subset Z of $\sqrt{-1}SM$ (resp., $\sqrt{-1}S^*M$) is also said to be **(properly) convex** if $\tau^{-1}(x) \cap Z$ (resp., $\pi^{-1}(x) \cap Z$) is so for each x in M . For a subset Z of $\sqrt{-1}SM$ its polar set Z° is, by definition, $\{(x, \sqrt{-1}\xi\infty) \in \sqrt{-1}S^*M | \langle v_x, \xi_x \rangle > 0 \text{ holds for each } x + \sqrt{-1}v_x 0 \text{ in } Z\}$. The polar set Z° of a subset Z of $\sqrt{-1}S^*M$ is defined in the same way. (vi) Let U be an open subset of $\sqrt{-1}SM$ such that $\tau^{-1}(x) \cap U$ is a nonvoid connected set for each x in M . Let V denote $U^{\circ\circ}$. Then we have (a) The restriction mapping $\rho: \Gamma(V; \tilde{\mathcal{A}}) \rightarrow \Gamma(U; \tilde{\mathcal{A}})$ is a bijection. Here $\Gamma(V; \tilde{\mathcal{A}})$, etc. denotes the space of global sections of $\tilde{\mathcal{A}}$ over V , etc. (b) $0 \rightarrow \tilde{\mathcal{A}}(U) \rightarrow \mathcal{B}(M) \xrightarrow{sp} \mathcal{C}(\sqrt{-1}S^*M - U^\circ)$ is an exact sequence. (vii) Let $f(x)$ be a hyperfunction on M . Then the following two statements are equivalent: (a) $sp(f)_{(x_0, \sqrt{-1}\xi_0\infty)} = 0$. (b) There exist a finite family of open subsets U_j of $\sqrt{-1}SM$ whose polar set U_j° does not contain $(x_0, \sqrt{-1}\xi_0\infty)$ and φ_j in $\Gamma(U_j, \tilde{\mathcal{A}})$ such that $f = \sum_j b(\varphi_j)$. Then we say that f is **micro-analytic** at $(x_0, \sqrt{-1}\xi_0\infty)$.

(3) Operations on Microfunctions. Let M and N be real analytic manifolds, and let f be a real analytic mapping from N to M . We denote by T_N^*M the kernel of the natural mapping from $N \times_M T^*M$ to T^*N . It is also called a **conormal bundle supported by N** . The associated sphere bundle is denoted by S_N^*M . Denote by ρ and ϖ the natural mappings from $N \times_M \sqrt{-1}S^*M \setminus \sqrt{-1}S_N^*M$ to $\sqrt{-1}S^*N$ and from $N \times_M \sqrt{-1}S^*M \setminus \sqrt{-1}S_N^*M$ to $\sqrt{-1}S^*M$, respectively. (i) Let \mathcal{B}_M^* denote the

sheaf $\{u \in \mathcal{B}_M \mid \text{S.S. } u \cap \sqrt{-1} S_M^* M = \emptyset\}$. Then we have the following two canonical homomorphisms: $f^*: \mathcal{B}_M^* \rightarrow \mathcal{B}_N^*$ and $f^*: \mu_! \varpi^{-1} \mathcal{C}_M \rightarrow \mathcal{C}_N$. Here and in what follows, for a continuous mapping φ from N to M and a sheaf \mathcal{F} on N , $\varphi_! \mathcal{F}$ denotes the sheaf on M defined by assigning $\{s \in \Gamma(\varphi^{-1}(U); \mathcal{F}) \mid \varphi|_{\text{supp } s}: \text{supp } s \rightarrow U \text{ is } \dagger\text{proper}\}$ to each open subset U of M . These two homomorphisms are consistent. We call each of them a **substitution** (homomorphism) and denote it by $(f^* u)(y) = u(f(y))$. (ii) Let v_M denote the sheaf of \dagger densities. Then there exist the following two canonical homomorphisms: $f_*: f_!(\mathcal{B}_N \otimes_{\mathcal{B}_N} v_N) \rightarrow \mathcal{B}_M \otimes_{\mathcal{B}_M} v_M$ and $f_*: \varpi_! \rho^{-1}(\mathcal{C}_N \otimes_{\mathcal{C}_N} v_N) \rightarrow \mathcal{C}_M \otimes_{\mathcal{C}_M} v_M$. These two homomorphisms are consistent. We call each an **integration along a fiber** and denote it by $(f_* u)(x) = \int_{f^{-1}(x)} u$. (iii) Using the result in (i), we can define the **product** $u_1 u_2$ of two hyperfunctions u_1 and u_2 , if $\text{S.S. } u_1 \cap (\text{S.S. } u_2)^a = \emptyset$. Furthermore, the singularity spectrum of $u_1 u_2$ is contained in $\{(x, \sqrt{-1}(\theta \xi_1 + (1 - \theta) \xi_2) \infty) \mid (x, \sqrt{-1} \xi_1 \infty) \in \text{S.S. } u_1, (x, \sqrt{-1} \xi_2 \infty) \in \text{S.S. } u_2, 0 \leq \theta \leq 1\} \cup \text{S.S. } u_1 \cup \text{S.S. } u_2$.

F. Microdifferential Operators [22]

(1) Microlocal Operators. Let M be a real analytic manifold, and define the sheaf \mathcal{L}_M on $\sqrt{-1} S_M^*(M \times M) = \sqrt{-1} S^* M$ by $\mathcal{H}_{\sqrt{-1} S_M^*(M \times M)}^0(\mathcal{C}_{M \times M} \otimes v_M)$. A section $K(x, y)$ dy of \mathcal{L}_M naturally determines an integral operator $\mathcal{K}: u(y) \rightarrow \int K(x, y) u(y) dy$. An operator thus obtained is called a **microlocal operator**, because it acts on the sheaf \mathcal{C}_M of microfunctions as a sheaf homomorphism. Usually we identify an operator \mathcal{K} and a kernel function $K(x, y) dy$. (i) \mathcal{L}_M is a sheaf of rings by the natural composition. The unit element of \mathcal{L}_M is $\delta(x - y) dy$. It acts on \mathcal{C}_M as an identity operator. (ii) Let $K(x, y) dy$ be the kernel function of a microlocal operator \mathcal{K} defined near $(x_0, \sqrt{-1} \xi_0 \infty)$. Then its **adjoint operator** \mathcal{K}^* is, by definition, the microlocal operator defined near $(x_0, -\sqrt{-1} \xi_0 \infty)$ with the kernel function $K(y, x) dy$. The operation $*$ is a sheaf isomorphism between \mathcal{L}_M and \mathcal{L}_M^a , where a denotes the antipodal mapping (\rightarrow Section E).

(2) Microdifferential Operators. A microdifferential operator is an analog of a microlocal operator in the complex domain. By a procedure similar to that used to define the sheaf of microfunctions we first define the sheaf $\mathcal{C}_{Y|X}^\infty$ of **holomorphic microfunctions** for a submanifold Y of X (\rightarrow [22, definition 1.1.7 on p. 319], where it is denoted by $\mathcal{C}_{Y|X}$). It follows from the definition that $\mathcal{C}_{Y|X}^\infty$ is sup-

ported by $P_Y^*(Y \times X)$, which is identified with $Y \times_X P^* X$. Here and in what follows $P^* X$, etc., denotes the cotangential projective bundle of X , etc. Then the sheaf \mathcal{E}_X^∞ of **microdifferential operators (of infinite order)** is, by definition, $\mathcal{C}_{X|X \otimes X}^\infty \otimes_{\mathcal{C}_X} \Omega_X$, where Ω_X denotes the sheaf of holomorphic $\dim X$ -forms.

Remark. Several notations are used to denote \mathcal{E}_X^∞ in the literature. For example, [22] uses the symbol \mathcal{D}_X and calls it the sheaf of **pseudodifferential operators**. As in the case of microfunctions, some authors use the symbol \mathcal{E}_X^∞ to denote a sheaf on $T^* X$. In this case $\mathcal{E}_X^\infty|_X$ is, by definition, \mathcal{D}_X^∞ , the sheaf of linear differential operators (of infinite order). One should be careful in these notational confusions in referring to papers using microdifferential operators. We note also that the symbols \mathcal{D} and \mathcal{E} have nothing to do with the symbols in distribution theory.

We now list the basic properties of microdifferential operators.

(i) When X is a complexification of a real analytic manifold M , $\mathcal{E}_X^\infty|_{\sqrt{-1} S^* M}$ is a subring of \mathcal{L}_M .

(ii) Let Ω be an open subset of $P^* X$. Using a local coordinate system (x) on X , we define $\hat{\Omega}$ by $\{(x, \xi) \in \mathbb{C}^n \times (\mathbb{C}^n - \{0\}) \mid (x, \xi \infty) \in \Omega\}$. Let $\{p_j(x, \xi)\}_{j \in \mathbb{Z}}$ be a sequence of holomorphic functions on $\hat{\Omega}$ satisfying the following conditions: (1) $p_j(x, \xi)$ is homogeneous of degree j in ξ . (2) For each $\varepsilon > 0$ and each compact subset K of $\hat{\Omega}$, there exists a constant $C_{\varepsilon, K}$ such that $\sup_K |p_j(x, \xi)| \leq C_{\varepsilon, K} \varepsilon^{|j|} / j!$ ($j \geq 0$) holds. (3) For each compact subset K of $\hat{\Omega}$, there exists a constant R_K such that $\sup_K |p_j(x, \xi)| \leq R_K^{-j} (-j)! (j < 0)$ holds. Then there is a one-to-one correspondence between the space of such sequences and the space of sections of \mathcal{E}_X^∞ over Ω .

(iii) A sequence satisfying the conditions in (ii) is called a **symbol sequence**, and the corresponding section of \mathcal{E}_X^∞ is denoted by $\sum_{j \in \mathbb{Z}} p_j(x, D_x)$. If we define a subsheaf $\mathcal{E}_X(m)$ of \mathcal{E}_X^∞ by $\{P = \sum_j p_j(x, D_x) \in \mathcal{E}_X^\infty \mid p_j(x, \xi) = 0 (j \geq m + 1)\}$, it is independent of the choice of the local coordinate systems. A microdifferential operator belonging to $\mathcal{E}_X(m)$ is said to be of **order** (at most) m . We denote $\bigcup_m \mathcal{E}_X(m)$ by \mathcal{E}_X and call a section of \mathcal{E}_X a **microdifferential operator of finite order**.

(iv) Let $\Phi_\lambda(z)$ denote $\Gamma(\lambda)/(-z)^\lambda$, where its branch is chosen so that $\Phi_\lambda(-1) = \Gamma(\lambda)$. When $\lambda = 0, -1, -2, \dots$, we consider its \dagger finite part. Let Ω be a complex neighborhood of $(x_0, \sqrt{-1} \xi_0 \infty) \in \sqrt{-1} S^* \mathbb{R}^n \cong \mathbb{R}^n \times \sqrt{-1} S^{n-1}$. Using a symbol sequence $\{p_j(z, \xi)\}$ on $\hat{\Omega}$, we define a multivalued holomorphic function $K(z, w, \zeta)$ by $\sum_j p_j(z, \zeta) \Phi_{n+j}(\langle z - w, \zeta \rangle)$ and consider its boundary value from the domain

$\text{Re}\langle z-w, \xi \rangle < 0$. We denote the resulting microfunction by

$$\sum_j p_j(x, \sqrt{-1} \xi) \Phi_{n+j}(\sqrt{-1}(\langle x-y, \xi \rangle + \sqrt{-1} 0)).$$

Then

$$K(x, y) = (2\pi)^{-n} \int \left\{ \sum_j p_j(x, \sqrt{-1} \xi) \Phi_{n+j}(\langle x-y, \sqrt{-1} \xi \rangle - 0) \right\} \omega(\xi)$$

is a well-defined microfunction in a neighborhood of $(x_0, x_0; \sqrt{-1}(\xi_0, -\xi_0)\infty)$ whose support is contained in the antidiagonal set $\Delta^a = \{(x, y; \sqrt{-1}(\xi, \eta)\infty) \in \sqrt{-1}S^*(M \times M) | x=y, \xi+\eta=0\}$. Here $\omega(\xi)$ is the volume element of the $(n-1)$ -dimensional sphere S^{n-1} . Hence $K(x, y)dy$ defines a microlocal operator. The mapping which associates $K(x, y)dy$ with $\{p_j(x, \xi)\}$ is compatible with the inclusion mapping stated in (i).

(v) By using the \dagger plane wave decomposition of the δ -function due to F. John (\rightarrow 125 Distributions and Hyperfunctions CC), we find that microdifferential operators are a natural generalization of linear differential operators: A linear differential operator corresponds to a symbol sequence $\{p_j(x, \xi)\}_{j \geq 0}$, where p_j is a polynomial with respect to ξ .

(vi) (a) Let $P = \sum_j p_j(x, D_x)$ and $Q = \sum_k q_k(x, D_x)$ be microdifferential operators. Then their composition $R = P \circ Q$ is a microdifferential operator with the symbol sequence $\{r_l\}_{l \in \mathbb{Z}}$ given by

$$r_l(x, \xi) = \sum_{i=j+k-|a|} \frac{1}{\alpha!} D_\xi^a p_j(x, \xi) D_x^a q_k(x, \xi).$$

Here $D_\xi^a = \partial^{|\alpha|} / \partial \xi_1^{\alpha_1} \dots \partial \xi_n^{\alpha_n}$ and $\alpha! = \alpha_1! \dots \alpha_n!$ for the \dagger multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$. (b) Let $P = \sum_j p_j(x, D_x)$ be a microdifferential operator. Let $\delta(x-y)$ denote the residue class $[1]$ of the left $\mathcal{E}_{X \times X}$ -Module $\mathcal{E}_{X \times X} / (\sum_{k=1}^n \mathcal{E}_{X \times X}(x_k - y_k) + \sum_{k=1}^n \mathcal{E}_{X \times X}(\partial/\partial x_k + \partial/\partial y_k))$. ("Module" means sheaf of modules.) Then there exists a unique microdifferential operator $R = \sum_l r_l(y, D_y)$ such that $P(x, D_x)\delta(x-y) = R(y, D_y)\delta(x-y)$. Furthermore, $r_l(x, \xi)$ is given by

$$\sum_{l=j-|a|} \frac{(-1)^j}{\alpha!} D_\xi^a D_x^a p_j(x, \xi).$$

R is called the **adjoint operator** of P and is denoted by P^* . When X is a complexification of the real manifold M , it coincides with the adjoint operator $P^* \in \mathcal{L}_M^a$.

(vii) For a microdifferential operator P in $\mathcal{E}_X(m)$, we define its **principal symbol** $\sigma_m(P)$ by $p_m(x, \xi)$. The principal symbol $\sigma_m(P)$ is in-

dependent of the choice of local coordinate system. It gives an isomorphism between $\mathcal{E}_X(m)/\mathcal{E}_X(m-1)$ and $\mathcal{O}_{T^*X}(m)$, the sheaf of holomorphic functions on T^*X which are homogeneous of degree m with respect to ξ .

(viii) (a) Let P be in $\mathcal{E}_X(m)_{(x_0, \xi_0)}$. Assume that $\sigma_m(P)(x_0, \xi_0) \neq 0$. Then its inverse P^{-1} (i.e., $PP^{-1} = P^{-1}P = 1$) exists in $\mathcal{E}_X(-m)_{(x_0, \xi_0)}$. (b) Let P and G be in $\mathcal{E}_X(m)_{(x_0, \xi_0)}$, $\mathcal{E}_X(l)_{(x_0, \xi_0)}$, respectively. Suppose that $H_{\sigma_f(G)}^j(\sigma_m(P))(x_0, \xi_0) = 0$ ($j=0, \dots, p-1$) and that $H_{\sigma_f(G)}^p(\sigma_m(P))(x_0, \xi_0) \neq 0$. (Here $H_f(g)$ is, by definition, the \dagger Poisson bracket $\{f, g\}$ of f and g (\rightarrow 82 Contact Transformations)). Then for each S in $\mathcal{E}_{X, (x_0, \xi_0)}$ we can find Q and R in $\mathcal{E}_{X, (x_0, \xi_0)}$ so that $S = QP +$

R with $(\text{ad } G)^p R \stackrel{\text{def}}{=} [\overbrace{G, \dots, G}^p, [G, R], \dots] = 0$ holds. This result is usually referred to as the **Sp ath-type division theorem (for microdifferential operators)**. In particular, when $G = x_n$ and $(x_0, \xi_0) = (0; 1, 0, \dots, 0)$, R has the form $\sum_{k=0}^{p-1} R^{(k)}(x, D') D_n^k$. Here $R^{(k)}(x, D') = R^{(k)}(x, D_1, \dots, D_{n-1})$. As a corollary to this expression we find the following (**Weierstrass-type preparation theorem (for microdifferential operators)**): Let P be as above, and let $G = x_n$. Then we can find Q and W in $\mathcal{E}_{X, (0; 1, 0, \dots, 0)}$ such that $P = QW$ with invertible Q and $W = D_n^p + \sum_{k=0}^{p-1} W^{(k)}(x, D') D_n^k$, where $W^{(k)}$ belongs to $\mathcal{E}_X(p-k)$ and $\sigma_{p-k}(W^{(k)})(0; 1, 0, \dots, 0) = 0$.

(ix) **Quantized contact transformation.** (a) Let X be an n -dimensional complex manifold and Ω an open subset of P^*X . Let P_j ($j=1, 2, \dots, n$) be in $\mathcal{E}_X(1)(\Omega)$ and Q_j ($j=1, \dots, n$) in $\mathcal{E}_X(0)(\Omega)$. Assume that $[P_j, P_k] = [Q_j, Q_k] = 0$ and $[P_j, Q_k] = \delta_{jk}$ hold ($1 \leq j, k \leq n$). Let φ be the contact transformation from Ω to $P^*\mathbb{C}^n$ defined by $p \mapsto (\sigma_0(Q_1)(p), \dots, \sigma_0(Q_n)(p), \sigma_1(P_1)(p), \dots, \sigma_1(P_n)(p))$. Then there exists a unique \mathbb{C} -algebra homomorphism $\Phi: \varphi^{-1}\mathcal{E}_{\mathbb{C}^n} \rightarrow \mathcal{E}_X|_\Omega$ such that $\Phi(x_j) = Q_j$ and $\Phi(D_j) = P_j$ ($j=1, \dots, n$). Furthermore, Φ is an isomorphism, $\Phi\mathcal{E}_{\mathbb{C}^n}(m) = \mathcal{E}_X(m)$ holds, and $\sigma_m(\Phi(R)) = \sigma_m(R) \circ \varphi$ holds for R in $\mathcal{E}_{\mathbb{C}^n}(m)$. We call the pair (φ, Φ) a **quantized contact transformation**. In the above situation, the $\mathcal{E}_{X \times \mathbb{C}^n}$ -Module

$$\mathcal{M} = \mathcal{E}_{X \times \mathbb{C}^n} / \left(\sum_{j=1}^n \mathcal{E}_{X \times \mathbb{C}^n}(x_j - Q_j) + \sum_{j=1}^n \mathcal{E}_{X \times \mathbb{C}^n}(-D_j - P_j) \right)$$

is a simple holonomic system (\rightarrow Section H) whose support is the graph of φ^a . Let u be the canonical generator of \mathcal{M} , i.e., the residue class of 1 in \mathcal{M} . Then $R^*u = \Phi(R)u$ holds for $R \in \mathcal{E}_{\mathbb{C}^n}$. (b) Conversely, let φ be a contact transformation from a neighborhood of p in P^*X to $P^*\mathbb{C}^n$, and let u be a generator of a simple holonomic system whose support is the graph

of φ^a . Then a \mathbb{C} -algebra isomorphism $\Phi: \varphi^{-1}\mathcal{E}_{\mathbb{C}^n} \rightarrow \mathcal{E}_X$ is defined in a neighborhood of p through $R^*u = \Phi(R)u$ ($R \in \mathcal{E}_{\mathbb{C}^n}$), and (φ, Φ) becomes a quantized contact transformation. (c) In particular, let p be a point in $\sqrt{-1} S^*M$ and Ω its complex neighborhood. Let (φ, Φ) be a real quantized contact transformation defined on Ω (i.e., φ maps $\sqrt{-1} S^*M$ to $\sqrt{-1} \cdot S^*\mathbf{R}^n$). Then, in a neighborhood of $(p, \varphi(p)^a) \in \sqrt{-1} S^*(M \times \mathbf{R}^n)$, there exists a microfunction solution $K(y, x) \neq 0$ of the equations $x_j K(y, x) = Q_j K(y, x)$, $-(\partial/\partial x_j) K(y, x) = P_j K(y, x)$ ($j = 1, \dots, n$) ($x \in \mathbf{R}^n, y \in M$). Such a microfunction is unique up to constant multiple. The integral operator $\mathcal{K}: v(x) \mapsto \int K(y, x)v(y)dy$ gives rise to a sheaf isomorphism between $\varphi^{-1}\mathcal{E}_{\mathbf{R}^n}$ and \mathcal{E}_M in a neighborhood of p . Furthermore, we have $\mathcal{K}(Rv) = \Phi(R)(\mathcal{K}v)$ for $v \in \mathcal{E}_{\mathbf{R}^n}$ and $R \in \mathcal{E}_{\mathbb{C}^n}$. This \mathcal{K} is the counterpart of the Fourier integral operator (\rightarrow Section C).

(x) Algebraic properties of \mathcal{E}_X^∞ and \mathcal{E}_X . (a) \mathcal{E}_X is \dagger coherent as a left \mathcal{E}_X -Module, and its stalk is a \dagger left Noetherian ring. (b) \mathcal{E}_X^∞ is \dagger faithfully flat over \mathcal{E}_X . (c) \mathcal{E}_X is \dagger flat over $\mathcal{E}_X(0)$. (d) \mathcal{E}_X is flat over $\pi^{-1}\mathcal{Q}_X$.

G. Microdifferential Equations [22, 24]

(1) **Background.** A system \mathcal{M} of **microdifferential equations** (of finite order) is by definition a \dagger coherent left (or right) \mathcal{E}_X -Module, i.e., there exists locally an exact sequence of the form $\mathcal{E}_X^p \rightarrow \mathcal{E}_X^q \rightarrow \mathcal{M} \rightarrow 0$. For a coherent \mathcal{E} -Module \mathcal{M} , the support $\text{Supp } \mathcal{M}$ of \mathcal{M} is an important geometric object associated with \mathcal{M} . It is called the **characteristic variety** of \mathcal{M} . For a coherent \mathcal{D} -Module \mathcal{M} , its characteristic variety is by definition $\text{Supp}(\mathcal{E} \otimes_{\mathcal{D}} \mathcal{M})$. It is often denoted by $\text{S.S. } \mathcal{M}$. Since a microdifferential operator is a microlocal operator, the result in (viii) (a) of Section F (3) asserts that $\mathcal{E}xt_{\mathcal{E}}^j(\mathcal{M}, \mathcal{E}_M)$ is supported by the characteristic variety of M intersected with $\sqrt{-1} S^*M$. Now, it is known that $V = \text{Supp } \mathcal{M}$ is \dagger involutive (= involutive, in involution) in P^*X , namely, $f|_V = g|_V = 0$ entails $\{f, g\}|_V = 0$ [22, theorem 5.3.2 on p. 453]. One of the most important problems in microlocal analysis is to study how much information V can give concerning the structure of \mathcal{M} itself, and hence that of $\mathcal{E}xt_{\mathcal{E}}^j(\mathcal{M}, \mathcal{E}_M)$. The epoch-making discovery of [22] is that V determines the structure of $\mathcal{E}^\infty \otimes_{\mathcal{E}} \mathcal{M}$ at generic points of V as follows.

(2) **Structure Theorems.** The fundamental result of [22, theorem 5.3.7 on p. 455] is the following theorem for coherent \mathcal{E} -Modules:

Structure theorem 1. Let \mathcal{M} be a coherent \mathcal{E}_X -Module satisfying the following conditions:

(1) $\mathcal{E}xt_{\mathcal{E}_X}^j(\mathcal{M}, \mathcal{E}_X) = 0$ for $j \neq d$. (2) $V \stackrel{\text{def}}{=} \text{Supp } \mathcal{M}$ is **regular** at $p \in V$ in the sense that V is non-singular near p and that $\omega|_V(p) \neq 0$ for the \dagger canonical 1-form ω . Then, through a quantized contact transformation (φ, Φ) , $\mathcal{E}^\infty \otimes \mathcal{M}$ is isomorphic to a direct summand of a direct sum of finite copies of **partial de Rham system** $\mathcal{N}_0 = \mathcal{E}^\infty_{\mathbb{C}^n} / (\sum_{j=1}^d \mathcal{E}^\infty_{\mathbb{C}^n} \partial/\partial z_j)$ with $\varphi(p) = (0; 0, \dots, 0, 1) \in P^*\mathbb{C}^n$.

By studying the canonical form of V under real contact transformations, [22] further gives the following structure theorem in a real domain, i.e., in $\sqrt{-1} S^*M$ for a real analytic manifold M .

Structure theorem 2. Let X be a complexification of a real analytic manifold M , and let \mathcal{M} be as in structure theorem 1. Let p be a point in $V \cap \sqrt{-1} S^*M$. Suppose that the following three conditions are satisfied: (3) $V \cap \bar{V}$ is regular at p . (4) $T_q(V) \cap T_q(\bar{V}) = T_q(V \cap \bar{V})$ holds for each q in $V \cap \bar{V}$. Here \bar{V} denotes the complex conjugate of V (with respect to $\sqrt{-1} S^*M$) and $T_q(V)$, etc., denotes the tangent space of V , etc., at q . (5) The generalized Levi form of V is of constant \dagger signature (a, b) near p , where the **generalized Levi form** of V is, by definition, the Hermitian form

$$L(v) = \sum_{j,k} \{p_j, \bar{p}_k\}(x, \sqrt{-1} \xi) v_j \bar{v}_k$$

for the p_j such that $V = \bigcap_j p_j^{-1}(0)$. Then $\mathcal{E}^\infty \otimes \mathcal{M}$ is isomorphic to a direct summand of a direct sum of finite copies of the system $\mathcal{E}^\infty \otimes_{\mathcal{E}} \mathcal{N}$ considered in a neighborhood of $(x, \sqrt{-1} \xi) = (0, \sqrt{-1}(0, \dots, 0, 1)) \in \sqrt{-1} S^*\mathbf{R}^n$, where \mathcal{N} is given by

$$\begin{aligned} \frac{\partial}{\partial x_j} f &= 0 \quad (j = 1, \dots, r), \\ \left(\frac{\partial}{\partial x_{r+2k-1}} + \sqrt{-1} \frac{\partial}{\partial x_{r+2k}} \right) f &= 0 \quad (k = 1, \dots, s), \\ \left(\frac{\partial}{\partial x_{r+2s+l}} + \sqrt{-1} x_{r+2s+l} \frac{\partial}{\partial x_n} \right) f &= 0 \\ &\quad (l = 1, \dots, a), \\ \left(\frac{\partial}{\partial x_{r+2s+l}} - \sqrt{-1} x_{r+2s+l} \frac{\partial}{\partial x_n} \right) f &= 0 \\ &\quad (l = a+1, \dots, a+b). \end{aligned}$$

Here, $r = 2 \text{codim } V - \text{codim}(V \cap \bar{V})$ and $s = \text{codim}(V \cap \bar{V}) - \text{codim } V - (a+b)$.

The first (resp. second) type of equation in the above are called (partial) **de Rham equations** (resp. (partial) **Cauchy-Riemann equations**). The third and the fourth are called **Lewy-Mizohata equations** (of type (a, b)) after these authors' pioneering works [25, 26]. Thus any system is seen to be microlocally isomorphic to a mixture of these three types of equations, generically speaking. As a corollary to

this, structure theorem 2 clarifies the structure of microfunction solutions of \mathcal{M} as follows:

Structure theorem 3. Let M, X, \mathcal{M}, V , and p be as in structure theorem 2. Then $\mathcal{E}x_{\mathcal{E}_X}^j(\mathcal{E}_X^\alpha \otimes \mathcal{M}, \mathcal{E}_M) = 0$ ($j \neq \alpha$) holds, and the remaining cohomology group $\mathcal{F} = \mathcal{E}x_{\mathcal{E}_X}^a(\mathcal{E}_X^\alpha \otimes \mathcal{M}, \mathcal{E}_M)$ has the following structure in a neighborhood U of p : There exists an s -dimensional complex manifold Y , a real analytic manifold N , and a † smooth mapping φ from $V \cap U \cap \sqrt{-1} S^*M$ to $Y \times \sqrt{-1} S^*N$ such that $\mathcal{F} = \varphi^{-1} \mathcal{G}$ for a sheaf \mathcal{G} on $Y \times \sqrt{-1} S^*N$ that is a direct summand of \mathcal{H}^N , with \mathcal{H} being the solution sheaf of the partial Cauchy-Riemann equations associated with Y .

H. Holonomic Systems

A coherent (left) \mathcal{E} -Module \mathcal{M} is called **holonomic** if $\text{Supp } \mathcal{M}$ is † Lagrangian. A coherent (left) \mathcal{D} -Module \mathcal{M} is called holonomic if $\mathcal{E} \otimes_{\mathcal{D}} \mathcal{M}$ is so. Even though the term "holonomic" is currently used, another term, "**maximally overdetermined**," is used to describe the same object in some of the literature, including [22]. The importance of such a system lies in the fact that the space of its microfunction solutions is finite-dimensional [27, 29]. In this sense it resembles an ordinary differential equation. A holonomic system, however, does not satisfy condition (2) of structure theorem 1 of Section G, and its structure is rather complicated. A result which corresponds to structure theorem 1 in Section G is the following: Let V be an involutory submanifold of T^*X , and let \mathcal{E}_V be the subring of \mathcal{E}_X generated by $\{P \in \mathcal{E}_X(1) \mid \sigma_1(P)|_V = 0\}$. Then a coherent \mathcal{E}_X -Module \mathcal{M} defined on an open subset Ω of T^*X is said to have **regular singularities** along V if for any point p of Ω , there exist a neighborhood U of p and an \mathcal{E}_V -sub-Module \mathcal{M}_0 of \mathcal{M} defined on U which is coherent over $\mathcal{E}(0)$, and which generates \mathcal{M} as an \mathcal{E}_X -Module. A holonomic system is said to have R.S., which is an abbreviation for regular singularities, if it has R.S. along its support. Then for an arbitrary holonomic \mathcal{E} -Module \mathcal{M} we can find a holonomic \mathcal{E} -Module \mathcal{M}_{reg} with R.S. such that $\mathcal{E}^\infty \otimes_{\mathcal{E}} \mathcal{M}_{\text{reg}} \cong \mathcal{E}^\infty \otimes_{\mathcal{E}} \mathcal{M}$ holds. See [28] for the proof of this striking result and related topics on **holonomic systems with regular singularities**.

An elementary class of holonomic systems is that of **simple holonomic systems**. A holonomic \mathcal{E} -Module \mathcal{M} is called simple if there exists a left Ideal \mathcal{I} such that $\mathcal{M} = \mathcal{E}/\mathcal{I}$ and that the symbol Ideal $\{\sigma(P) \mid P \in \mathcal{I}\}$ coincides with the defining Ideal of $\text{Supp } \mathcal{M}$. Let u denote the generator $1 \bmod \mathcal{I}$ of a simple holonomic

system \mathcal{E}/\mathcal{I} . Suppose that $\Lambda = \text{Supp } \mathcal{M}$ is nonsingular. Then the **principal symbol** $\sigma_\Lambda(u)$ is defined as follows: For $P = \sum_j p_j(x, D_x) \in \mathcal{I}(m)$, let L_p denote the first-order linear differential operator

$$L_p = H_{p_m(x, \xi)} + \left(p_{m-1}(x, \xi) - \frac{1}{2} \sum_i \frac{\partial^2 p_m(x, \xi)}{\partial x_i \partial \xi_i} \right).$$

Here, H_{p_m} denotes the Hamiltonian vector field defined by $\{p_m, \cdot\}$. Let Ω_Λ and Ω_X denote the sheaf of n -forms of Λ and X , respectively, and let $\Omega_\Lambda^{\otimes 1/2}$ (resp. $\Omega_\Lambda^{\otimes 1/2} \otimes \Omega_X^{\otimes -1/2}$) denote a line bundle L such that $L^{\otimes 2}$ is isomorphic to Ω_Λ (resp. $\Omega_\Lambda \otimes \Omega_X^{\otimes -1}$). Since the line bundles $\Omega_\Lambda^{\otimes 1/2}$ and $\Omega_\Lambda^{\otimes 1/2} \otimes \Omega_X^{\otimes -1/2}$ do not exist globally in general, all the equations among their sections should be understood up to a constant multiplicative factor. When Λ is a purely imaginary Lagrangian submanifold of $\sqrt{-1} S^*M$ for a real analytic manifold M , these line bundles can be constructed globally by using the Keller-Maslov index (\rightarrow Section C). Let v and ψ denote respectively the first term and the second term in the right-hand side of the above definition of L_p . Then $\tilde{L}_p: \Omega_\Lambda^{\otimes 1/2} \rightarrow \Omega_\Lambda^{\otimes 1/2}$ is given by $\tilde{L}_p(s) = (1/2s)L_v(s^2) + \psi s$ for $s \in \Omega_\Lambda^{\otimes 1/2}$, where $L_v(s^2)$ denotes the † Lie derivative of s^2 along v . One can then prove that the system of equations $\tilde{L}_p s = 0$ ($P \in \mathcal{I}$) admits locally one and exactly one nonzero solution s in $\Omega_\Lambda^{\otimes 1/2}$ up to a constant factor. Then the principal symbol $\sigma_\Lambda(u)$ of u equals $s \otimes \sqrt{dx} \in \Omega_\Lambda^{\otimes 1/2} \otimes \Omega_X^{\otimes -1/2}$ by definition. The principal symbol $\sigma_\Lambda(u)$ is homogeneous with respect to ξ , and its homogeneous degree is called the **order** of u and is denoted by $\text{ord}_\Lambda(u)$. The microlocal structure of a simple holonomic \mathcal{E} -Module with a nonsingular characteristic variety is determined by the order of its generator as follows: (a) Let $\mathcal{E}u$ and $\mathcal{E}v$ denote simple holonomic \mathcal{E} -Modules with the same characteristic variety Λ . Then $\mathcal{E}u$ is isomorphic to $\mathcal{E}v$ if and only if $\text{ord}_\Lambda(u) - \text{ord}_\Lambda(v)$ is an integer. (b) Let $\mathcal{E}u$ be a simple holonomic system, and let α denote the order of u . Then through a suitable quantized contact transformation, $\mathcal{E}u$ is isomorphic to $\mathcal{E}_{C^n} w$, where w satisfies

$$\left(z_1 \frac{\partial}{\partial z_1} + \left(\alpha + \frac{1}{2} \right) \right) w = 0,$$

$$\frac{\partial w}{\partial z_j} = 0 \quad (j = 2, \dots, n).$$

Thus the microlocal structure of a simple holonomic system $\mathcal{M} = \mathcal{E}u$ is fairly simple at nonsingular points of its characteristic variety. Moreover a Hartogs-type theorem for micro-differential equations [28, ch. I, §2] entails that if $\text{Supp } \mathcal{M}$ has the form $\Lambda_1 \cup \Lambda_2$ with Lagrangian manifolds Λ_1 and Λ_2 such that

Microlocal Analysis

$\text{codim}_{\Lambda_1}(\Lambda_1 \cap \Lambda_2) \geq 2$; then \mathcal{M} has the form $\mathcal{M}_1 \oplus \mathcal{M}_2$ with $\text{supp } \mathcal{M}_j = \Lambda_j$ ($j = 1, 2$). Hence the case where $\text{codim}_{\Lambda_1}(\Lambda_1 \cap \Lambda_2) = 1$ is important. Suppose Λ_1 and Λ_2 are nonsingular in this case and $T_x \Lambda_1 \neq T_x \Lambda_2$ at a point $x \in \Lambda_1 \cap \Lambda_2$. Then, through a quantized contact transformation (ψ, Φ) , the system \mathcal{M} is isomorphic to \mathcal{E}_{C^*w} , with w satisfying the following equations defined near $(0; dz_1, \infty)$:

$$\left(z_1 \frac{\partial}{\partial z_1} + \left(\alpha_1 + \frac{1}{2} \right) \right) w = 0,$$

$$\left(z_2 \frac{\partial}{\partial z_2} + \left(\alpha_1 - \alpha_2 + \frac{1}{2} \right) \right) w = 0,$$

$$\frac{\partial w}{\partial z_j} = 0 \quad (j = 3, \dots, n),$$

where $\alpha_j = \text{ord}_{\Lambda_j}(u)$ ($j = 1, 2$) with $\psi(\Lambda_1) = \{z_1 = 0, \xi_2 = \dots = \xi_n = 0\}$ and $\psi(\Lambda_2) = \{z_1 = z_2 = 0, \xi_3 = \dots = \xi_n = 0\}$. For more general cases and an application \rightarrow [30].

The theory of holonomic systems and its applications are being studied most intensively, raising the hope of establishing a unified theory of special functions of several variables; \rightarrow [23] and references cited therein.

I. History

Microlocal analysis means local analysis on the cotangent bundle. It emphasizes the importance of localization in cotangent bundles in analysis, which was pointed out by S. Mizohata [5, 6] immediately after the advent of singular integral operators in the works of A. P. Calderón and A. Zygmund [3, 4]. Since then, localization in the cotangent bundle has been used frequently in the theory of linear partial differential equations. R. T. Seeley [31] proved that the symbol of a singular integral operator is well defined on the cotangent bundle. Works by J. J. Kohn and L. Nirenberg [32] and L. Hörmander (*Comm. Pure Appl. Math.*, 18 (1965)) strengthened the trend of localizing the problem on the cotangent bundle. Although it seems that the term "microlocal analysis" first appeared in the literature in 1973 (T. Kawai, *Astérisque*, 2 and 3 (1973)), the basic part of the theory had been constructed during the period from 1969 to 1972 by M. Sato (*Proc. Intern. Conf. Functional Anal. and Related Topics*, 1969), Yu. V. Egorov [12], Hörmander [8, 20], J. J. Duistermaat and Hörmander (*Acta Math.*, 128 (1972)), M. Kashiwara and Kawai (*Proc. Japan Acad.*, 46 (1970)), and Sato, Kawai, and Kashiwara [22]. Apparently the work of V. P. Maslov [11] had an important influence on the work of Egorov. The most important part of Sato's

contribution was the fact that, through the construction of the sheaf of microfunctions he found that the singularities of hyperfunctions can be canonically dispersed over the cotangent bundle (\rightarrow Section E) and that a hyperfunction solution u of a linear differential equation $Pu = 0$ is concentrated on the characteristic variety when its singularities are thus dispersed (\rightarrow Section G). The last-stated fact was also formulated by Hörmander [8, 20] in the framework of distribution theory. The most important part of the contribution of Egorov [12] was the discovery that one can use an integral transformation introduced by V. I. Eskin [7] to find a transformation of pseudodifferential operators compatible with a homogeneous canonical transformation, i.e., a contact transformation, so that the commutation relations and the orders of the operators can be preserved. Hörmander (*Acta Math.*, 121 (1969)) independently introduced integral transformations of the same type, calling them Fourier integral operators. Egorov [13] and L. Nirenberg and F. Trèves [33] successfully used the transformation of operators to study the regularity and existence of solutions. Subsequently, Hörmander [8] elaborated the theory of Fourier integral operators. Kashiwara and Kawai (*Proc. Japan Acad.*, 46 (1970)) observed that a pseudodifferential operator in their sense (now called a microdifferential operator; \rightarrow Section F) gives rise to a sheaf homomorphism on the sheaf of microfunctions and that the structure of the microfunction solutions of pseudodifferential equations is determined by the principal symbol of the operator in question if it has simple characteristics. Then Sato, Kawai, and Kashiwara [22] succeeded in amalgamating these two theories, namely, the theory of microfunctions and the theory of the transformation of operators. (They called the transformation a quantized contact transformation in [22]). Such an amalgamation was also done independently by Hörmander [8, 20], who introduced the notion of the wave front set for distributions as a substitute for the support of microfunctions. Incidentally, it is noteworthy that C. Chandler and H. P. Stapp (*J. Math. Phys.*, 10 (1969)) and D. Iagolnitzer and Stapp (*Comm. Math. Phys.*, 14 (1969)) obtained a notion similar to the singularity spectrum in a physical context. Their results were later elaborated (around 1971–1973) by J. Bros and Iagolnitzer [19]. With the aid of the above-mentioned amalgamation of the theories, the works of Sato, Kawai, and Kashiwara [22], Hörmander [20], Duistermaat and Hörmander (*Acta Math.*, 121 (1969)), Kawai (*Publ. Res. Inst. Math. Sci.*, 7 (1971–1972)) and K. G. Andersson (*Trans. Amer. Math. Soc.*, 177 (1973)) have clarified the

importance of the bicharacteristic strip, which is a submanifold of a cotangent bundle, not a base manifold, as a carrier of singularities of solutions of pseudodifferential equations (\rightarrow Section G); and the works of Sato, Kawai, and Kashiwara [22], Sato (*Actes Congr. Internat. Math. Nice*, 1971) and Kawai (*Publ. Res. Inst. Math. Sci.*, 7 (1971–1972)) revealed the hidden mechanism of the celebrated counterexample of H. Lewy [25]. Among these, the contribution of Sato, Kawai, and Kashiwara [22] was most decisive and fundamental in that it first clarified the structure of a general system of pseudodifferential equations at generic points of the characteristic variety and then derived from it the above-quoted results on the structure of the solutions (\rightarrow Section G).

Microlocal analysis has now become one of the most important and basic concepts in the theory of linear partial differential equations and theoretical physics. For recent developments \rightarrow Hörmander [34] and V. Guillemin, Kashiwara, and Kawai [23] and references cited therein.

References

- [1] J. J. Duistermaat, Fourier integral operators, Lecture notes, Courant Inst. Math. Sci., 1973.
- [2] V. Guillemin and S. Sternberg, Geometric asymptotics, Amer. Math. Soc. Math. Surveys 14 (1977).
- [3] A. P. Calderón and A. Zygmund, Singular integral operators and differential equations, Amer. J. Math., 79 (1957), 901–921.
- [4] A. P. Calderón, Uniqueness in the Cauchy problem for partial differential equations, Amer. J. Math., 80 (1958), 16–36.
- [5] S. Mizohata, Systèmes hyperboliques, J. Math. Soc. Japan, 11 (1959), 205–233.
- [6] S. Mizohata, Note sur le traitement par les opérateurs d'intégrale singulière de Cauchy, J. Math. Soc. Japan, 11 (1959), 234–240.
- [7] G. I. Eskin, The Cauchy problem for hyperbolic systems in convolutions, Math. USSR-Sb., 3 (1967), 243–277.
- [8] L. Hörmander, Fourier integral operators I, Acta Math., 127 (1971), 79–183.
- [9] J. Chazarain (ed.), Fourier integral operators and partial differential equations, Lecture notes in math. 459, Springer, 1975.
- [10] J. B. Keller, Corrected Bohr Sommerfeld conditions for nonseparable systems, Ann. Phys., 4 (1958), 180–188.
- [11] V. P. Maslov, Théorie des perturbations et méthodes asymptotiques, Gauthier-Villars, 1972. (Original in Russian, 1965.)
- [12] Yu. V. Egorov, On canonical transformation of pseudodifferential operators (in Russian), Uspekhi Mat. Nauk, 25 (1969), 235–236.
- [13] Yu. V. Egorov, Conditions for solvability of pseudodifferential equations, Soviet Math. Dokl., 10 (1969), 1020–1022.
- [14] D. Fujiwara, On the boundedness of integral operators with highly oscillatory kernels, Proc. Japan Acad., 51 (1975), 96–99.
- [15] H. Kumano-go, A calculus of Fourier integral operators on \mathbf{R}^n and the fundamental solution for an operator of hyperbolic type, Comm. Partial Diff. Eq., 1 (1976), 1–44.
- [16] K. Asada and D. Fujiwara, On some oscillatory integral transformations in $L^2(\mathbf{R}^n)$, Japan. J. Math., 4 (1978), 299–361.
- [17] A. Sommerfeld, Optics, Lectures on Theoretical Physics, vol. 4, Academic Press, 1964.
- [18] P. D. Lax, Asymptotic solution of oscillatory initial value problems, Duke Math. J., 24 (1957), 627–646.
- [19] D. Iagolnitzer, Analytic structure of distributions and essential support theory, Structural Analysis and Collision Amplitude, North-Holland, 1976, 295–358.
- [20] L. Hörmander, Uniqueness theorem and wave front sets for solution of linear differential equation with analytic coefficients, Comm. Pure Appl. Math., 24 (1971), 671–704.
- [21] J. M. Bony, Equivalence des diverses notions de spectre singulier analytique, Séminaire Goulaouic-Schwartz 1976–77, École Polytechnique.
- [22] M. Sato, T. Kawai, and M. Kashiwara, Microfunctions and pseudodifferential equations, Lecture notes in math. 287, Springer, 1973, 265–529.
- [23] V. Guillemin, M. Kashiwara, and T. Kawai, Seminar on microlocal analysis, Ann. math. studies 93, Princeton Univ. Press, 1979.
- [24] J. E. Björk, Rings of differential operators, North-Holland, 1979.
- [25] H. Lewy, An example of a smooth linear partial differential equation without solution, Ann. Math., 66 (1957), 155–158.
- [26] S. Mizohata, Solutions nulles et solutions nonanalytiques, J. Math. Kyoto Univ., 1 (1962), 271–302.
- [27] M. Kashiwara, On the maximally overdetermined system of linear differential equations I, Publ. Res. Inst. Math. Sci., 10 (1974–1975), 563–579.
- [28] M. Kashiwara and T. Kawai, On holonomic systems of microdifferential equations III, Systems with regular singularities, Publ. Res. Inst. Math. Sci., 17 (1981), 813–979.
- [29] M. Kashiwara and T. Kawai, Finiteness theorem for holonomic systems of microdifferential equations, Proc. Japan Acad., 52 (1976), 341–343.
- [30] M. Sato, M. Kashiwara, T. Kimura, and

Minimal Submanifolds

T. Oshima, Microlocal analysis of prehomogeneous vector spaces, *Inventiones Math.*, 62 (1980), 117–179.

[31] R. T. Seeley, Singular integrals on compact manifolds, *Amer. J. Math.*, 81 (1959), 658–690.

[32] J. J. Kohn and L. Nirenberg, An algebra of pseudodifferential operators, *Comm. Pure Appl. Math.*, 18 (1965), 269–305.

[33] L. Nirenberg and F. Trèves, On local solvability of linear partial differential equations II, *Comm. Pure Appl. Math.*, 23 (1970), 459–510.

[34] L. Hörmander, Seminar on singularities of solutions of linear partial differential equations, *Ann. math. studies* 91, Princeton Univ. Press, 1979.

275 (VII.14) Minimal Submanifolds

A. General Remarks

An immersion f of an m -dimensional manifold M with boundary ∂M (possibly empty) into a Riemannian manifold N is called **minimal** if the †mean curvature vector field H of M with respect to the induced Riemannian metric vanishes identically. Then M is called a **minimal submanifold** of N . This definition comes from the following variational problem: By a smooth †variation of f is meant a smooth mapping $F: I \times M \rightarrow N$, where $I = (-1, 1)$, such that each $f_t = F(t, \cdot): M \rightarrow N$ is an immersion, $f_0 = f$, and $f_t|_{\partial M} = f|_{\partial M}$ for all $t \in I$. Let dV_t be the volume element of the metric induced by f_t , and set $V(t) = \int_M dV_t$, the volume of M at time t . Then the first variation of the volume is expressed as

$$\left. \frac{dV}{dt} \right|_{t=0} = -m \int_M \left\langle H, F_* \frac{\partial}{\partial t} \right|_{t=0} \rangle dV_0,$$

where $\partial/\partial t$ denotes the canonical vector field along the I factor in $I \times M$. Thus the mean curvature vector field H of f vanishes identically if and only if $dV/dt|_{t=0} = 0$ for all variations of f . Therefore a minimal submanifold gives an extremal of the volume integral, though neither necessarily minimal nor of the least volume.

In the case $N = \mathbf{R}^n$, an immersion $x: M \rightarrow \mathbf{R}^n$ is viewed as a vector-valued function, and the mean curvature vector field H is expressed as $H = \Delta x/m$, where Δ denotes the †Laplace-Beltrami operator $-g^{ij} \nabla_i \nabla_j$. Thus x is minimal if and only if each component function of x is harmonic. In particular, there is no compact minimal submanifold without boundary in \mathbf{R}^n .

The history of the theory of minimal submanifolds goes back to J. L. Lagrange, who studied minimal surfaces in the 3-dimensional Euclidean space \mathbf{R}^3 . In 1762 he developed his algorithm for the †calculus of variations, which can be applied to higher-dimensional problems and is now known as the †Euler-Lagrange equation. For instance, let D be a domain in the plane \mathbf{R}^2 and $z = f(x, y)$, $(x, y) \in D$, the equation of a surface in \mathbf{R}^3 . As a necessary condition for the surface to have the least area among the surfaces with fixed boundary, Lagrange obtained a †quasilinear elliptic partial differential equation of the second order, called the **minimal surface equation**:

$$(1 + z_y^2)z_{xx} - 2z_x z_y z_{xy} + (1 + z_x^2)z_{yy} = 0.$$

Before this, in 1744, L. Euler had found that a †catenoid is a minimal surface. In 1766, J. B. M. C. Meusnier showed that a right †helicoid is a minimal surface. Besides catenoids and helicoids, in 1834, H. F. Scherk found that the surface defined by $z = \log(\cos y) - \log(\cos x)$ is a minimal surface, which is called **Scherk's surface**.

In the latter half of the 19th century, †Plateau's problem (\rightarrow Section C; 334 Plateau's Problem) was studied extensively by O. Bonnet, B. Riemann, K. Weierstrass, A. Enneper, G. Darboux, and others. The problem is stated as follows: Given a †Jordan curve Γ in \mathbf{R}^3 (or in \mathbf{R}^n), find a surface of least area having Γ as its boundary. On the other hand, in 1866, Weierstrass gave a general formula, called the **Weierstrass-Enneper formula** (\rightarrow Section B (5)), to express a simply connected minimal surface in terms of a complex analytic function and a meromorphic function with certain properties. The formula allows one to construct a great variety of minimal surfaces by choosing those functions.

The existence of a minimal surface of disk type having a prescribed boundary curve was first obtained in 1930 by J. Douglas and T. Radó independently as a solution to Plateau's problem, admitting singularities. The result was improved by R. Courant for the case of finitely many boundary curves by the method of †Dirichlet integrals. The method was carried out further by C. B. Morrey for the generalized Plateau's problem in a Riemannian manifold (\rightarrow Section C (5)). Indeed, the Euclidean space was replaced by any complete Riemannian manifold which is metrically well behaved at infinity. For example, any compact or any †homogeneous Riemannian manifold is in this class.

The existence proof of minimal surfaces cannot in general be applied directly to the case of higher-dimensional minimal submanifolds. The notion of varifolds (\rightarrow Section G (2))

and the corresponding generalization of a minimal submanifold, called a minimal variety (\rightarrow Section G (2)), were then introduced by F. Almgren, who proved the existence of a minimal variety in a compact Riemannian manifold. It was also proved that a minimal variety is approximated by regular minimal submanifolds. This work has been developed as geometric measure theory by E. R. Reifenberg, W. H. Fleming, H. Federer, F. Almgren, and others (\rightarrow Section G).

The study of minimal surfaces given as the graph of a real-valued function of two variables leads to that of solutions of the minimal surface equation. In 1915, S. Bernstein proved that a minimal surface $z = f(x, y)$ defined on the entire plane \mathbf{R}^2 must be a plane. Subsequently, this was generalized to the following problem: Is a minimal hypersurface $x_{n+1} = f(x_1, \dots, x_n)$ in \mathbf{R}^{n+1} defined on the entire space \mathbf{R}^n a hyperplane? The answer turns out to be affirmative for $n \leq 7$ and negative for $n \geq 8$ (\rightarrow Section F (1)).

In general, when a bounded domain D in \mathbf{R}^n and a continuous function φ on its boundary ∂D are given, the problem of finding a minimal hypersurface M defined by the graph of a real-valued function f on \bar{D} with $f|_{\partial D} = \varphi$ gives rise to a typical \dagger Dirichlet problem. The basic questions are those of existence, uniqueness, and regularity of solutions. These were first studied by Radó for $n = 2$ and later by L. Bers, R. Finn, H. Jenkins, J. Serrin, R. Osserman, and others (\rightarrow Section D (1)).

Minimal submanifolds of a sphere have interesting properties; some are analogous to those in Euclidean space but some are not. Among them, J. Simons, in 1968, gave a differential equation that the norm of the second fundamental form of a minimal submanifold in a sphere should satisfy. He then showed that a totally geodesic submanifold is isolated among the compact minimal submanifolds in the sphere. S. S. Chern, M. P. do Carmo, S. Kobayashi, N. Ejiri, T. Ito, T. Otsuki, N. Wallach, and others have made further contributions to this subject (\rightarrow Section F (2)).

B. Minimal Surfaces

(1) **Branched minimal surfaces.** A minimal surface M in \mathbf{R}^n is an immersed surface with vanishing mean curvature vector field. M equipped with an atlas of \dagger isothermal coordinates is viewed as a \dagger Riemann surface. A **branch point** of a \dagger harmonic mapping $f: M \rightarrow \mathbf{R}^n$ is a point at which the differential df_p is zero. A harmonic mapping $f: M \rightarrow \mathbf{R}^n$ of a Riemann surface M is called a **branched** (or

generalized) **minimal immersion** if it is conformal except at the branch points, and the image $f(M)$ is then called a **branched minimal surface**. The solution to Plateau's problem given by Douglas and Radó is a branched minimal surface.

(2) **Maximum principle.** When $n = 3$, the following maximum principle for minimal surfaces holds: If M_1 and M_2 are two connected branched minimal surfaces in \mathbf{R}^3 such that for a point $p \in M_1 \cap M_2$, the surface M_1 locally lies on one side of M_2 near p , then M_1 and M_2 coincide near p .

(3) **Convex hull property.** In general, every branched minimal surface with boundary in \mathbf{R}^n lies in the **convex hull** of its boundary curve, the smallest closed convex set containing the boundary.

(4) **Reflection principle.** If the boundary curve of a branched minimal surface contains a straight line γ , then the surface can be analytically continued as a branched minimal surface by reflection across γ . Based on this principle, the following holds: Let Γ be an analytic Jordan curve in \mathbf{R}^n and $f: M \rightarrow \mathbf{R}^n$ a branched minimal immersion with boundary Γ . Then f is analytic up to the boundary, i.e., $f(M)$ is contained in the interior of a larger branched minimal surface (H. Lewy). The smooth version of this theorem was given by S. Hildebrandt: If $f: M \rightarrow \mathbf{R}^n$ is a branched minimal immersion with smooth boundary curve, then f is smooth up to the boundary.

(5) **Weierstrass-Enneper formula.** Every simply connected minimal surface in \mathbf{R}^3 is represented in the form

$$x_k(\zeta) = \operatorname{Re} \left\{ \int_0^\zeta \phi_k(z) dz \right\} + c_k, \quad k = 1, 2, 3,$$

where $\phi_1 = f(1 - g^2)/2$, $\phi_2 = \sqrt{-1} f(1 + g^2)/2$, $\phi_3 = fg$, and c_k is a constant. Here, $g(\zeta)$ is a meromorphic function on a domain D in the complex ζ -plane, and $f(\zeta)$ is an analytic function on D with the property that at each point ζ , where $g(\zeta)$ has a pole of order m , $f(\zeta)$ has a zero of order $2m$, D being either the unit disk or the entire plane. This formula is quite useful for constructing various minimal surfaces. For instance, on setting $f \equiv 1$ and $g(\zeta) = \zeta$, **Enneper's surface** is obtained:

$$(x_1, x_2, x_3) = \left(u - \frac{u^3}{3} + uv^2, v - \frac{v^3}{3} + vu^2, u^2 - v^2\right), \\ (u, v) \in \mathbf{R}^2.$$

Another feature of the formula is that general theorems about minimal surfaces can be obtained by translating statements about analytic functions into the corresponding minimal-surface ones. For example, a surface $f: M \rightarrow \mathbf{R}^3$ is minimal if and only if the \dagger Gauss mapping

$G: M \rightarrow S^2$ is **antiholomorphic**, i.e., the complex conjugate of the mapping is holomorphic, when S^2 is viewed as the Riemann sphere with complex structure by the stereographic projection (from the south pole) onto the complex plane. As a consequence, a complete minimal surface in \mathbf{R}^3 is a plane or else the normals to the surface are everywhere dense in S^2 (R. Osserman).

(6) **Stability.** A minimal submanifold M is called **stable** if for every compact region on M all the second variations of the volume are positive. For instance, if $M \rightarrow \mathbf{R}^n$ is a minimal hypersurface and $f\nu$ is a variation vector field, ν being the unit normal vector field of M and f a smooth function with compact support on M , then the second variation formula for the volume $V(t)$ is given by

$$\left. \frac{d^2 V}{dt^2} \right|_{t=0} = \int_M (|\nabla f|^2 - |A|^2 f^2) dv_0,$$

where $|A|$ denotes the norm of the second fundamental form of M . In particular, when $n=3$, then $-|A|^2/2=K$, the Gauss curvature of M . Therefore a minimal surface M in \mathbf{R}^3 is stable if and only if

$$\int_M (|\nabla f|^2 + 2Kf^2) dV > 0$$

for any smooth function f with compact support on M . It follows that the stability of minimal surface M in \mathbf{R}^3 is related to the boundary value problem of a linear elliptic operator $L = \Delta - 2K$, Δ being the Laplacian on M .

Namely, a minimal surface M in \mathbf{R}^3 is stable if and only if the first eigenvalue $\lambda_1(D)$ of L on any bounded domain D in M is nonnegative.

In connection with the \dagger Gauss mapping, a sufficient condition for a domain D in a minimal surface to be stable was obtained by H. Schwarz in 1885: If a minimal surface M in \mathbf{R}^3 has one-to-one Gauss mapping $G: M \rightarrow S^2$, then a domain $D \subset M$ such that $G(D)$ is contained in a hemisphere of S^2 is stable. This was generalized by J. L. Barbosa and M. do Carmo as follows: If the area $A(G(D))$ of the spherical image $G(D)$ is less than 2π , then D is stable.

This result is sharp in the sense that for any $\varepsilon > 0$ there exists an unstable domain D with $A(G(D)) = 2\pi + \varepsilon$. As an analog to Bernshtein's theorem, the following holds: A complete and stable minimal surface in \mathbf{R}^3 is a plane. H. Mori has also obtained results in this regard.

As for the existence of unstable minimal surfaces, it is known that if there are two distinct stable minimal surfaces with the same smooth boundary curve Γ in \mathbf{R}^3 , then there exists an unstable branched minimal surface with Γ as its boundary (M. Morse and C. Tompkins, M. Schiffman).

C. Plateau's Problem (\rightarrow 334 Plateau's Problem)

Let Γ be a \dagger rectifiable \dagger Jordan curve in \mathbf{R}^n and $D = \{(x, y) \in \mathbf{R}^2 \mid x^2 + y^2 < 1\}$. Then there exists a continuous mapping $f: \bar{D} \rightarrow \mathbf{R}^n$ such that (a) $f|_{\partial D}$ maps homeomorphically onto Γ ; (b) $f|_D$ is harmonic and **almost conformal**, i.e., $\langle f_x, f_y \rangle = 0$ and $|f_x| = |f_y|$ in D with $|df| > 0$ except at isolated branch points; (c) the induced area of f is the least among the class of piecewise smooth surfaces bounding Γ with (a). This mapping f is called the **classical solution** or the **Douglas-Radó solution** to Plateau's problem for Γ , which may have singularities called \dagger branch points. The resulting surface S is a branched minimal disk. This theorem establishes the existence of a surface of least area among all surfaces homeomorphic to a disk. It has been generalized by R. Courant for Γ consisting of finitely many rectifiable Jordan curves.

For a branched minimal disk S bounded by smooth Γ in \mathbf{R}^3 , there is a relation, called the **Gauss-Bonnet-Sasaki-Nitsche formula**, among the \dagger total curvature $\kappa(\Gamma)$ of Γ , the total curvature of S , and the orders of branch points:

$$1 + \sum (m_x - 1) + \sum M_\beta + \frac{1}{2\pi} \int_S |K| dA \leq \frac{1}{2\pi} \kappa(\Gamma),$$

where K denotes the Gauss curvature of S , $m_x - 1$ the orders of the interior branch points, and $2M_\beta$ the orders of the boundary branch points, which must be even.

(1) **Regularity.** A minimal disk of least area in \mathbf{R}^3 has no boundary branch points when Γ is real analytic (R. Gulliver and F. Leslie) or when Γ is smooth and the total curvature $\kappa(\Gamma)$ is less than 4π (J. C. C. Nitsche). In general, a classical solution for smooth Γ in \mathbf{R}^3 cannot have infinitely many branch points. A remarkable fact is that every classical solution to Plateau's problem in \mathbf{R}^3 is free of branch points in its interior, i.e., is a regular immersion (R. Osserman).

(2) **Embeddedness.** A classical solution is not necessarily an embedding; it may have self-intersections. For instance, if Γ is knotted in \mathbf{R}^3 , then every solution must have self-intersections. It is known, however, that immersed minimal disks of least area in \mathbf{R}^3 which can self-intersect only in their interiors are embeddings. In particular, if Γ is an **extremal** Jordan curve, i.e., if Γ lies on the boundary of its \dagger convex hull, then the classical solution for Γ is an embedding (F. Tomi and A. J. Tromba; W. H. Meeks and S. T. Yau). If the topological type is not specified, then there always exists an embedded minimal surface. That is, if Γ is the union of any finite collection of disjoint

smooth Jordan curves in \mathbf{R}^3 , then there exists a compact embedded minimal surface with boundary Γ which is smooth up to the boundary (R. Hardt and L. Simon). However, there exists an unknotted Jordan curve that never bounds an embedded minimal disk.

(3) **Uniqueness.** In general, the classical solutions to Plateau's problem do not have the uniqueness property. Radó was the first who gave a condition on a boundary curve to guarantee the uniqueness of the minimal disk. Namely, if a Jordan curve Γ in \mathbf{R}^n admits a one-to-one orthogonal projection onto a convex curve in a plane \mathbf{R}^2 in \mathbf{R}^n , then the classical solution to Plateau's problem for Γ is free of branch points and can be expressed as the graph over this plane. When $n=3$, the solution is unique. Another geometric condition on Γ has been given by Nitsche: An analytic Jordan curve Γ in \mathbf{R}^3 with total curvature $\kappa(\Gamma) \leq 4\pi$ (or a smooth Γ with $\kappa(\Gamma) < 4\pi$) bounds a unique immersed minimal disk. Moreover, the generic uniqueness holds: In the space \mathcal{A} of all smooth Jordan curves in \mathbf{R}^n with suitable topology, there exists an open and dense subset \mathcal{B} such that for any Γ in \mathcal{B} there exists a unique area-minimizing minimal disk (F. Morgan and A. J. Tromba).

(4) **Finiteness.** As for the finiteness of the classical solutions to Plateau's problem, several conditions on boundary curves are known. An analytic Jordan curve in \mathbf{R}^3 bounds only finitely many minimal disks of least area (F. Tomi). An analytic extremal Jordan curve in \mathbf{R}^3 bounds only finitely many minimal disks with relative minima of area (H. Beeson). A smooth Jordan curve Γ with total curvature $\kappa(\Gamma) \leq 6\pi$ bounds only finitely many minimal disks (Nitsche). Moreover, generically, i.e., for an open and dense subset of boundary curves, there are at most finitely many minimal surfaces with given boundary, relative minima or not (R. Böhme and Tromba).

(5) **Generalized Plateau's problem.** C. B. Morrey's setting of the generalized Plateau's problem is as follows: A homotopically trivial rectifiable Jordan curve Γ is given in an n -dimensional Riemannian manifold N . Let D denote a disk in \mathbf{R}^2 . Find a mapping $f: \bar{D} \rightarrow N$ such that (a) $f|_{\partial D}$ maps homeomorphically onto Γ , (b) the induced area of f is the least among the class of piecewise smooth surfaces in N bounded by Γ with (a). Obviously when $N = \mathbf{R}^n$, this is the classical Plateau's problem. A solution was given by Morrey under the assumption that N is **homogeneously regular**; i.e., that there exist $0 < k < K$ such that for any point $y \in N$ there is a local coordinate system (V, Φ) around y for which $\Phi(V) = \{x \in \mathbf{R}^n | \|x\| < 1\}$ and the Riemannian metric

$\sum g_{\alpha\beta} dx^\alpha dx^\beta$ satisfies

$$k \sum (v^\alpha)^2 \leq \sum g_{\alpha\beta} v^\alpha v^\beta \leq K \sum (v^\alpha)^2$$

for any x and $(v_1, \dots, v_n) \in \mathbf{R}^n$. Any compact or any homogeneous Riemannian manifold is homogeneously regular. Morrey's solution is as follows: If N is a homogeneously regular Riemannian manifold and if Γ is a homotopically trivial rectifiable Jordan curve in N , then there exists a branched minimal immersion $f: D \rightarrow N$ with least area bounded by Γ with (a).

The regularity of Morrey's solution is similar to the classical one. If Γ is smooth, then so is the solution f up to the boundary. If, furthermore, $\dim N = 3$, then the solution f is an immersion in its interior. If N and Γ are real analytic and $\dim N = 3$, then the solution f is an immersion up to the boundary.

D. Existence Problems of Minimal Submanifolds

(1) **The Dirichlet problem for the minimal surface equation.** When the graph of a (vector-valued) function is a minimal submanifold in some Euclidean space, the function must satisfy the so-called †minimal surface equation. To be precise, let D be a (bounded) domain in \mathbf{R}^n , $f: D \rightarrow \mathbf{R}^k$ a (vector-valued) function, and put $F: D \rightarrow \mathbf{R}^{n+k}$: $F(x) = (x, f(x))$ for $x \in D$. Let M be the graph of f : $M = F(D)$. Then M is minimal if and only if f (or F) satisfies one of the following equations:

if $n=2$ and $k=1$,

$$(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy} = 0; \quad (1)$$

if n is arbitrary and $k=1$,

$$\sum_{i=1}^n \frac{\partial}{\partial x^i} \left(\frac{1}{W} \frac{\partial f}{\partial x^i} \right) = 0 \quad \text{where} \quad W = \sqrt{1 + |\nabla f|^2}; \quad (2)$$

if $n=2$ and k is arbitrary,

$$(1 + |f_y|^2)f_{xx} - 2(f_x \cdot f_y)f_{xy} + (1 + |f_x|^2)f_{yy} = 0; \quad (3)$$

and if n and k are arbitrary,

$$\sum_{i,j=1}^n g^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j} = 0 \quad \text{where} \quad g_{ij} = \frac{\partial F}{\partial x^i} \cdot \frac{\partial F}{\partial x^j}, \quad (4)$$

and (g^{ij}) is the inverse matrix of (g_{ij}) .

The basic problem for this class of equations is the †Dirichlet problem. For $n=2$ and $k=1$, the following theorem of Radó and Finn is fundamental: There exists a solution f of the Dirichlet problem corresponding to an arbitrary continuous function φ on the boundary ∂D of D if and only if D is convex. Since the difference of any two solutions of eq. (1) satisfies the maximum principle, there can be at most one solution f of the Dirichlet problem

corresponding to a given continuous function φ on ∂D . As for the removability of isolated singularities, the following is known: Every solution of eq. (1) in $0 < x^2 + y^2 < 1$ extends continuously to the origin. The extended function is smooth at the origin and satisfies eq. (1) in the full disk $x^2 + y^2 < 1$ (Finn).

For arbitrary n and $k = 1$, namely, for the case of minimal hypersurfaces, the following is known: Let D be a bounded domain in \mathbf{R}^n with smooth boundary. The Dirichlet problem for eq. (2) has a solution for any continuous function φ on ∂D if and only if the mean curvature of ∂D with respect to the inner normal is non-negative at every point. As for the uniqueness and regularity, the same results as above hold (Jenkins and Serrin).

For $n = 2$ and arbitrary k , namely, for the case of minimal surfaces in \mathbf{R}^{2+k} , exactly the same result of existence as in the classical case $k = 1$ holds. A solution exists for an arbitrary continuous vector-valued function on ∂D if and only if D is convex (Osserman). Though the removability of singularities holds under suitable restrictions, the uniqueness fails in this case.

For the last case, $n > 2$ and $k > 1$, there are essentially no results on either existence or uniqueness for the Dirichlet problem.

(2) **Existence of minimal surfaces and topology.** For a compact Riemannian manifold N , there are results on the existence of minimal surfaces related to the homotopy groups of N . Let f be a continuous mapping from a \dagger Riemann surface Σ_g of \dagger genus g into N . If the \dagger induced mapping $f_\# : \pi_1(\Sigma_g) \rightarrow \pi_1(N)$ of f on the \dagger fundamental groups is injective, then there exists a branched minimal immersion $h : \Sigma_g \rightarrow N$ such that $h_\# = f_\#$ on $\pi_1(\Sigma_g)$ and the induced area of h is the least among all mappings with the same action on $\pi_1(\Sigma_g)$. If furthermore $\pi_2(N) = 0$, then h can be deformed from f continuously (R. Schoen and S. T. Yau). If $\pi_2(N) \neq 0$, then there exists a generating set for $\pi_2(N)$ consisting of branched minimal immersions of spheres that minimize energy and area in their homotopy classes (J. Sacks and K. Uhlenbeck). If, furthermore, $\dim N = 3$, then the above minimal immersion h in the homotopy class is an embedding or a two-to-one covering mapping. In the second case, its image is an embedded real projective plane. If h' is another such least-area mapping, then either h' is equal to h up to a conformal reparametrization or the images of h and h' are disjoint (Meeks and Yau).

E. Gauss Mapping of Minimal Surfaces

On a connected, orientable surface M there exist local isothermal coordinates (x, y) at each

point. On putting $z = x + \sqrt{-1} y$, the metric has the form $ds^2 = 2F|dz|^2$, and M is viewed as a Riemann surface. A conformal immersion $f : M \rightarrow \mathbf{R}^n$ is minimal if and only if $\Delta f = 0$, or equivalently $\partial_z \partial_{\bar{z}} f = 0$. In the case $n = 3$, the \dagger Gauss mapping $G : M \rightarrow S^2$ of a surface $f : M \rightarrow \mathbf{R}^3$ is defined by assigning to each point $p \in M$ the unit normal vector translated parallel to the origin of \mathbf{R}^3 . A surface $f : M \rightarrow \mathbf{R}^3$ is minimal if and only if the Gauss mapping is antiholomorphic. In the general case, the Gauss mapping G is defined to be a mapping assigning to each point $p \in M$ the oriented tangent space $f_* T_p(M) \subset \mathbf{R}^n$. Thus G is a mapping from M into the \dagger Grassmann manifold $\tilde{M}_{n,2}(\mathbf{R}) = SO(n)/SO(2) \times SO(n-2)$ of oriented planes in \mathbf{R}^n , which is naturally diffeomorphic to the \dagger complex quadric Q_{n-2} in the complex projective space \mathbf{CP}^{n-1} . An immersion $f : M \rightarrow \mathbf{R}^n$ is minimal if and only if its Gauss mapping is antiholomorphic.

Let $f : M \rightarrow \mathbf{R}^n$ be a complete orientable minimal surface. Let χ be the Euler characteristic and $-\pi C = \int_M K dA$, the total curvature of $f(M)$. Assume that the total curvature is finite. Then the following results of Chern and Osserman are fundamental: (a) M is conformally a compact Riemann surface \bar{M} with finite number, say r , of points deleted; (b) C is an even integer and satisfies $C \geq 2(r - \chi) = 4g + 4r - 4$, where g is the genus of M ($=$ genus of \bar{M}); (c) if $f(M)$ does not lie in any proper affine subspace of \mathbf{R}^n , then $C \geq 4g + r + n - 3 \geq 4g + n - 2 \geq n - 2$ (F. Gackstatter); (d) if $f(M)$ is simply connected and nondegenerate, i.e., its image under the Gauss mapping does not lie in a hyperplane of \mathbf{CP}^{n-1} , then $C \geq 2n - 2$ and this inequality is sharp; (e) when $n = 3$, C is a multiple of 4, with the minimum value 4 attained only for Enneper's surface and the catenoid; (f) the Gauss mapping G of M extends to a mapping of \bar{M} whose image $G(\bar{M})$ is an algebraic curve in \mathbf{CP}^{n-1} lying in Q_{n-2} ; the total curvature of $f(M)$ is equal in absolute value to the area of $G(\bar{M})$, counting multiplicity; (g) $G(M)$ intersects a fixed number of times, say m (counting multiplicity), every hyperplane in \mathbf{CP}^{n-1} except for those hyperplanes containing any of the finite number of points of $G(\bar{M} - M)$; the total curvature of $f(M)$ equals $-2\pi m$.

In particular, if a complete minimal surface in \mathbf{R}^n has k \dagger ends, then the total curvature never exceeds $2\pi(\chi - k)$. Enneper's surface and the catenoid are the only two complete minimal surfaces in \mathbf{R}^3 whose Gauss mapping is one-to-one.

If the Gauss mapping of a complete minimal surface of finite total curvature in \mathbf{R}^3 omits more than 3 points of S^2 , then it is a plane (Osserman). F. Xavier proved that the Gauss mapping of any complete nonflat minimal

surface in \mathbf{R}^3 can omit at most 6 points of S^2 . It is an open question whether this can be improved to 4 points. The Gauss mapping of Scherk's surface omits exactly 4 points of S^2 .

F. Minimal Submanifolds

(1) **Bernshtein's problem.** As was mentioned in Section A, a minimal surface in \mathbf{R}^3 represented as the graph of a function on the entire plane \mathbf{R}^2 is a plane. This is known as **Bernshtein's theorem** (1915). This is, however, not true in \mathbf{R}^4 , namely, there is a minimal surface defined as the graph of a vector-valued function on the entire plane, which is not a plane. The **generalized Bernshtein problem** is stated as follows: Let $f: \mathbf{R}^n \rightarrow \mathbf{R}$ be a function satisfying the minimal hypersurface equation

$$\sum_{i=1}^n \frac{\partial}{\partial x^i} \left(\frac{1}{W} \frac{\partial f}{\partial x^i} \right) = 0,$$

where $W = \sqrt{1 + |\nabla f|^2}$. Is the graph of f an affine hyperplane? The answer is affirmative for $n \leq 7$ (E. De Giorgi, F. J. Almgren, J. Simons). For $n \geq 8$, it is negative (E. Bombieri, De Giorgi, E. Giusti). Though the problem turns out to be negative for $n \geq 8$, no concrete counterexample is known, even for $n = 8$.

(2) **Minimal submanifolds in the sphere.** Minimal submanifolds in the unit sphere S^n have some special properties. For example, there is no compact minimal submanifold of S^n contained in an open hemisphere. On the other hand, there is no stable minimal closed submanifold in S^n (B. Lawson, Simons). More generally, any closed minimal hypersurface in a Riemannian manifold with positive Ricci curvature is unstable (Simons). As for the existence of minimal surfaces in a sphere, the following theorem of Lawson is general: Any closed surface of any genus, except the real projective plane, can be minimally immersed into the unit sphere S^3 . As an analog to Bernshtein's theorem, a sphere S^2 , which is minimally immersed in S^3 , is necessarily totally geodesic. On the other hand, by making use of higher-order second fundamental forms and the relations among them, a complete description of the minimal immersions of \mathbf{R}^2 into S^n has been given by K. Kenmotsu.

As a general theorem which gives a necessary and sufficient condition for a Riemannian manifold to be minimally immersed in the unit sphere S^n , the following is fundamental: An isometric immersion x of an m -dimensional Riemannian manifold M into S^n , viewed as a vector-valued function into \mathbf{R}^{n+1} , is minimal if and only if $\Delta x = mx$ (T. Takahashi). As its immediate application, any m -dimensional compact homogeneous space whose linear

isotropy group is irreducible can be minimally immersed into the n -sphere of curvature λ/m , corresponding to any eigenvalue ($\neq 0$) of the Laplacian, where $n+1$ is the dimension of the eigenspace corresponding to λ (W. Y. Hsiang, Takahashi).

Since for the k -symmetric spaces of rank 1 the eigenvalues and the corresponding eigenspaces of the Laplacian are known, the minimal immersions of such spaces into the unit sphere have been determined as well as the rigidity of such minimal immersions (N. Wallach). For example, if the m -sphere of constant curvature c is minimally immersed into the unit n -sphere, but not contained in any great hypersphere, then for each nonnegative integer k , $c = m/k(k+m-1)$ and $n+1 \leq (2k+m-1) \cdot [(k+m-2)!/k!(m-1)!]$. The immersion is rigid if and only if $m=2$ or $k \leq 3$. Similar results have been known for the projective spaces over real numbers, complexes, quaternions, or Cayley numbers.

Simons showed that the scalar curvature ρ of a compact minimal submanifold M of dimension m in the unit $(m+p)$ -sphere is not greater than $m(m-1)$. Furthermore, if $\rho \geq m(m-1) - mp/(2p-1)$, then either $\rho = m(m-1)$ or $\rho = m(m-1) - mp/(2p-1)$. In the former case, M is totally geodesic and therefore is the unit sphere S^m . In the latter case, M is either a hypersurface of the unit $(m+1)$ -sphere which is isometric to the product $S^k(\sqrt{k/m}) \times S^{m-k}(\sqrt{(m-k)/m})$ of spheres of radius $\sqrt{k/m}$ and $\sqrt{(m-k)/m}$, respectively, called the **generalized Clifford torus** or the **Veronese surface** in S^4 . Here the **Clifford torus** is the torus $S^1(\sqrt{1/2}) \times S^1(\sqrt{1/2}) \subset S^3 \subset \mathbf{R}^4$, and the Veronese surface is defined as follows: Let (x, y, z) be the natural coordinate system in \mathbf{R}^3 and $(u^1, u^2, u^3, u^4, u^5)$ that in \mathbf{R}^5 . The mapping defined by

$$\begin{aligned} u^1 &= yz/\sqrt{3}, & u^2 &= zx/\sqrt{3}, & u^3 &= xy/\sqrt{3}, \\ u^4 &= (x^2 - y^2)/2\sqrt{3}, & u^5 &= (x^2 + y^2 - 2z^2)/6 \end{aligned}$$

gives an isometric immersion of $S^2(\sqrt{3})$ into S^4 . Two points (x, y, z) and $(-x, -y, -z)$ of $S^2(\sqrt{3})$ are mapped into the same point of S^4 , and thus the mapping defines an embedding of the real projective plane \mathbf{RP}^2 into S^4 . This embedded real projective plane in S^4 is called the Veronese surface.

T. Otsuki proved the following: Let M be a complete minimal hypersurface immersed in the unit $(n+1)$ -sphere with two principal curvatures. If their multiplicities are m and $n-m \geq 2$, then M is congruent to the generalized Clifford torus $S^m(\sqrt{m/n}) \times S^{n-m}(\sqrt{(n-m)/n}) \subset S^{n+1} \subset \mathbf{R}^{n+2}$. If one of their multiplicities is 1, then M is a hypersurface of S^{n+1} in $\mathbf{R}^{n+2} = \mathbf{R}^n \times \mathbf{R}^2$ whose orthogonal projection into \mathbf{R}^2

Minimal Submanifolds

is a curve of which the †support function $x(t)$ is a solution of the following nonlinear differential equation of the second order:

$$nx(1-x^2)(d^2x/dt^2) + (dx/dt)^2 + (1-x^2)(nx^2-1) = 0.$$

Furthermore, there are countably many compact minimal hypersurfaces immersed but not embedded in S^{n+1} . Only $S^{n-1}(\sqrt{(n-1)/n}) \times S^1(\sqrt{1/n})$ is minimally embedded in S^{n+1} . This just corresponds to the trivial solution $x(t) \equiv 1/\sqrt{n}$ of the above equation.

(3) **Minimal submanifolds in Riemannian manifolds.** In general, it is difficult to determine all the complete stable minimal submanifolds in a given Riemannian manifold. If, however, some curvature conditions are given to the Riemannian manifold, then a classification has been given: Let N be a complete 3-dimensional Riemannian manifold with non-negative scalar curvature ρ , and let M be a complete, stable, and orientable minimal surface in N .

(a) If M is compact, then either M is conformally equivalent to the †Riemann sphere S^2 or else it is a totally geodesic flat torus. Furthermore, if $\rho > 0$, then the latter case never occurs.

(b) If M is not compact, then M is conformally equivalent to the complex plane \mathbb{C} or a cylinder (D. Fischer Colbrie; R. Schoen).

As a generalization of the procedure for generating periodic minimal surfaces in \mathbb{R}^3 with octahedral or tetrahedral symmetry (Schwarz, 1867), T. Nagano and B. Smyth gave a construction procedure to generate periodic minimal surfaces in \mathbb{R}^n or n -tori with symmetry corresponding to any †Weyl group of rank n .

G. Minimal Varieties

Recent progress in the study of Plateau's problem in higher dimensions is closely related to the point of view of geometric measure theory and †the calculus of variations.

(1) **Integral currents.** Let T be a †current of degree k , or simply a k -current, defined on an open set U of \mathbb{R}^n , and let ∂T be the boundary of T , the exterior derivative of T . For simplicity, in the following only currents with compact support are considered.

The **mass** $M(T)$ of a k -current T is the dual norm $M(T) = \sup\{T(\varphi) \mid M(\varphi) \leq 1\}$ of the **comass** $M(\varphi)$ on k -forms, which is defined by $M(\varphi) = \sup\{\|\varphi(x)\| \mid x \in U\}$ with $\|\varphi(x)\| = \sup\{\langle \varphi(x), v_1 \wedge \dots \wedge v_k \rangle \mid v_1, \dots, v_k \text{ are orthonormal vectors at } x \in U\}$. It follows that $M(T)$ coincides with k -dimensional volume when T

is a k -current defined by integration on a k -dimensional simplex. We say that T is **normal** if $M(T)$ and $M(\partial T)$ are finite. Normal currents form a †Banach space with norm $N(T) = M(T) + M(\partial T)$.

A current T is called **rectifiable** if it can be approximated in the mass norm by currents of type $f_*\gamma$, where γ is a finite polyhedral chain with integer coefficients, $f: \text{supp } \gamma \rightarrow U$ is a †Lipshitz mapping of the support of γ into U , and $f_*\gamma$ is the current defined by means of $f_*\gamma(\varphi) = \gamma(f^*\varphi)$. If both T and ∂T are rectifiable, T is called an **integral current**. Integral currents give the appropriate notion of generalized manifolds with boundary in the study of higher-dimensional Plateau problems.

One of the fundamental properties of integral currents is the compactness, stated as follows: Given a compact subset $K \subset U$ and a number $c > 0$, the set of integral currents T such that $\text{supp } T \subset K$ and $N(T) \leq c$ is sequentially compact in the †weak topology. Since the mass $M(T)$ is †lower semicontinuous in the weak topology, it follows that Plateau's problem can always be solved in the space of integral currents (\rightarrow Section C; 334 Plateau's Problem). The question arises: To what extent is this a satisfactory solution to the problem? For codimension 1, it is known from the work of Federer and others that an integral $(n-1)$ -current of least mass in \mathbb{R}^n is nonsingular in codimension < 7 . In particular, in \mathbb{R}^n , $n \leq 7$, every integral $(n-1)$ -current of least mass is an analytic manifold. In general codimensions, it is known that the set of regular points is dense (Reifenberg, Almgren, and others).

The definition of rectifiable and integral currents carries over to those on a Riemannian manifold M in a natural way. Then the space of integral currents $I_*(M)$ on M with the boundary operator ∂ , for which $\partial^2 = 0$, form a chain complex. It is then a fundamental theorem, due to Federer and Fleming, in homological integration theory that there is a natural isomorphism $H_*(I_*(M)) \cong H_*(M; \mathbb{Z})$ of the homology of the complex of integral currents with the integral singular homology groups of M . From this it follows that if M is a compact Riemannian manifold M , then each class $\alpha \in H_p(I_*(M)) \cong H_p(M; \mathbb{Z})$ contains an integral current of least mass among all integral currents in α .

For a homology class of codimension 1, it has also been proved by Almgren that if M is a compact Riemannian manifold of dimension ≤ 7 , then for each $\alpha \in H_{n-1}(M; \mathbb{Z})$ there exists a finite collection of mutually disjoint, compact oriented minimal hypersurfaces S_1, \dots, S_l embedded in M and integers n_1, \dots, n_l such that the integral current $\sum_{j=1}^l n_j S_j$ represents α . On the other hand, it has been recently proved

that every compact Riemannian manifold of dimension ≤ 6 contains a nonempty closed embedded minimal hypersurface (J. T. Pitts).

(2) **Varifolds.** The theory of integral currents developed by Federer, Fleming, and others yields reasonable spaces for purposes of the calculus of variations. However, it is not entirely feasible to use these in the study of the actual soap films that occur in physical experiments. It turns out that we should work in a more set-theoretic fashion and give up the notion of orientability and the boundary operator. A convenient theory for describing these physical experiments is the theory of varifolds developed by Almgren.

A k -dimensional **varifold** on a Riemannian manifold M is a \dagger Radon measure on the bundle $G_k(M)$ of unoriented tangent k -planes of M . For simplicity, only varifolds with compact support are considered. These are regarded as continuous linear functionals on the space $C(G_k(M))$ of continuous functions on $G_k(M)$. Thus a k -dimensional submanifold S of finite volume embedded in M determines a varifold V_S by integration:

$$V_S(f) = \int_S f(T_x S) d\mathcal{H}^k(x), \quad f \in C(G_k(M)),$$

where \mathcal{H}^k denotes the k -dimensional \dagger Hausdorff measure on M . More generally, to any rectifiable current there corresponds an underlying varifold obtained by neglecting orientations.

Let $V_k(M)$ denote the space of k -dimensional varifolds on M . Given a varifold $V \in V_k(M)$, the **mass** $M(V)$ is defined to be the V -measure of the total space, i.e., $M(V) = V(1)$. Let $f: M \rightarrow M$ be a C^1 -mapping. Then f induces naturally a mapping f_* of $V_k(M)$ into itself. In particular, if X is a smooth vector field on M with associated flow f_t , then there is a smooth function $M(t) = M(f_{t*} V)$. A varifold $V \in V_k(M)$ is then called a k -dimensional **minimal variety** in M if the first variation $(d/dt)M(t)|_{t=0} = 0$ for all smooth vector fields on M . For example, minimal submanifolds, complex analytic subvarieties of Kähler manifolds and integral currents of least mass are minimal varieties.

An appropriate notion of rectifiable and integral varifolds is defined, and an analog of the compactness theorem can be obtained. As a consequence, it was proved by Almgren using \dagger Morse theory methods that if M is an n -dimensional compact Riemannian manifold, then for each p , $1 \leq p \leq n-1$, there exists at least one minimal variety of dimension p . As for the regularity of minimal varieties, it is known that if V is a k -dimensional variety, then in the support of V there is a relatively open dense subset that is a regular minimal submanifold of dimension k (W. K. Allard).

H. Maximal Hypersurfaces in Minkowski Space

Let L^{n+1} be a Minkowski space, i.e., $L^{n+1} = \{(x_1, \dots, x_n, t) | (x_1, \dots, x_n) \in \mathbf{R}^n, t \in \mathbf{R}\}$ with the \dagger Lorentzian metric $\sum_{i=1}^n (dx_i)^2 - (dt)^2$. Let M be a \dagger spacelike hypersurface of L^{n+1} so that the induced metric on M is Riemannian. If the mean curvature vector field H of M , defined in the same way as in the Riemannian case, vanishes identically, then M is called **maximal**. In contrast to the Riemannian case, by the variation in the normal direction of H , the volume increases, provided $H \neq 0$. However, the equation describing a maximal hypersurface is similar to that for the Riemannian case (\rightarrow Sections D (1), F (1)). Indeed, let D be a domain in \mathbf{R}^n and $f: D \rightarrow \mathbf{R}$ a smooth function. Then the graph of f in L^{n+1} is a maximal hypersurface if and only if f satisfies the following quasilinear partial differential equation of the second order:

$$(1 - |\nabla f|^2) \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} + \sum_{i,j=1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \frac{\partial^2 f}{\partial x_i \partial x_j} = 0.$$

This is \dagger elliptic for $|\nabla f| < 1$, since M is spacelike. E. Calabi proposed a problem similar to Bernshtein's in 1968. Contrary to Bernshtein's case, the answer is affirmative for any n .

Namely, a maximal hypersurface in L^{n+1} defined as the graph of a function on the entire space \mathbf{R}^n is a hyperplane in \mathbf{R}^{n+1} . More generally, a maximal hypersurface, which is a closed set in L^{n+1} , is a hyperplane (S. Y. Cheng and S. T. Yau). It is also known that any maximal hypersurface in L^{n+1} is stable.

References

- [1] F. J. Almgren, Plateau's problem, An invitation to varifold geometry, Benjamin, 1966.
- [2] R. Courant, Dirichlet's principle, conformal mapping, and minimal surfaces, Interscience, 1950.
- [3] H. Federer, Geometric measure theory, Springer, 1969.
- [4] H. Federer, Colloquium lectures on geometric measure theory, Bull. Amer. Math. Soc., 84 (1978) 291–338.
- [5] D. Hoffman and R. Osserman, The geometry of the generalized Gauss map, Mem. Amer. Math. Soc., 236 (1980).
- [6] H. B. Lawson, Minimal varieties, Proc. Symp. Pure Math., 27 (1975), 143–175.
- [7] H. B. Lawson, Lectures on minimal submanifolds I, Publish or Perish, 1980.
- [8] C. B. Morrey, Multiple integrals in the calculus of variations, Springer, 1966.

- [9] J. C. C. Nitsche, On new results in the theory of minimal surfaces, *Bull. Amer. Math. Soc.*, 71 (1965), 195–270.
- [10] J. C. C. Nitsche, *Vorlesungen über Minimalflächen*, Springer, 1975.
- [11] R. Osserman, Minimal varieties, *Bull. Amer. Math. Soc.*, 75 (1969), 1092–1120.
- [12] R. Osserman, A survey of minimal surfaces, Van Nostrand, 1969.
- [13] T. Radó, On the problem of Plateau, Springer, 1933.
- [14] A. J. Tromba, On the number of simply connected minimal surfaces spanning a given curve, *Mem. Amer. Math. Soc.*, 194 (1977).
- [15] Minimal submanifolds and geodesics, *Proc. Japan-US Seminar*, Tokyo, 1977, North-Holland, 1979.
- [16] S. Nishikawa, On maximal spacelike hypersurfaces in a Lorentzian manifold, *Nagoya Math. J.*, 95 (1984), 117–124.
- [17] O. Kobayashi, Maximal surfaces with conelike singularities, *J. Math. Soc. Japan*, 36 (1984), 609–617.

276 (I.6) Model Theory

A. Language

Every mathematical theory has an appropriate language. To determine a language for a theory means to determine a language for the related mathematical system. Such a language consists of the following symbols (the symbols given here are examples of only one notational system).

- (1) Symbols that express logical concepts (*logical symbols): $\forall, \exists, \neg, \wedge, \vee, \rightarrow$;
- (2) *free variables: a_0, a_1, a_2, \dots ;
- (3) *bound variables: x_0, x_1, x_2, \dots ;
- (4) symbols that denote individual objects (individual constants): $c_0, c_1, c_2, \dots, c_\alpha, \dots$;
- (5) *function symbols: $f_0, f_1, f_2, \dots, f_\alpha, \dots$;
- (6) *predicate symbols: $P_0, P_1, P_2, \dots, P_\alpha, \dots$.

The *cardinalities of the sets of symbols in (4), (5), and (6) are arbitrary, except that there must be at least one predicate symbol. It is assumed that each set of symbols is *well ordered. Also, it is understood that to each f_j in (5) there corresponds a positive integer i_j , while to each P_j there corresponds a nonnegative integer (these integers are called the number of arguments of f_j and P_j , respectively).

In practice, other kinds of languages are also dealt with. One example is a system with infinitely long expressions, which permits

*transfinite ordinals for the numbers of arguments of f_j and P_j , and which has the extended concepts $\bigwedge_{\alpha < \beta}, \bigvee_{\alpha < \beta}, \exists x_1 \dots \exists x_\alpha \dots$ for $\alpha < \beta$, and $\forall x_1 \forall x_2 \dots \forall x_\alpha \dots$ for $\alpha < \beta$, where α and β are transfinite ordinals. Another language includes variables of higher *type as well as \forall and \exists over those variables. Free and bound variables may not be distinguished typographically. In that case a variable not bound by \forall or \exists in a *formula is called **free**. (The notion of a formula is defined later.) To simplify this discussion, however, we restrict ourselves to the *first-order predicate language with a typographic distinction between free and bound variables. We also assume that there are only a countable number of variables, and hence we use natural numbers as subscripts.

Set $L_1 = \{\text{logical symbols}\}$, $L_2 = \{a_0, a_1, a_2, \dots\}$, $L_3 = \{x_0, x_1, x_2, \dots\}$, $L_4 = \{c_0, c_1, c_2, \dots\}$, $L_5 = \{f_0, f_1, f_2, \dots\}$, $L_6 = \{P_0, P_1, P_2, \dots\}$, $L = \langle L_1, L_2, L_3, L_4, L_5, L_6 \rangle$. To determine a language L is to specify such a list L . Since $\forall, \exists, \neg, \wedge, \vee, \rightarrow$, are normally used for L_1 , a_0, a_1, a_2, \dots for L_2 , and x_0, x_1, x_2, \dots for L_3 , we may assume that these are fixed, and hence to determine a language is to determine $\langle L_4, L_5, L_6 \rangle$. We take $\langle L_1, L_2, L_3 \rangle$ just described and assume an arbitrary but fixed $L = \langle L_4, L_5, L_6 \rangle$. First we define the notions term of L (or L -term) and formula of L (or L -formula).

Definition of the **terms** of L (L -terms): (1) Each free variable a_j is an L -term. (2) Each individual constant c_j of L is an L -term. (3) If f_j is a function symbol of L , i_j is the number of arguments of f_j , and each of t_1, \dots, t_{i_j} is an L -term, then $f_j(t_1, \dots, t_{i_j})$ is also an L -term. (4) The L -terms are only those constructed by (1)–(3).

A term that does not contain a free variable is called a **closed term**.

Definition of the **formulas** of L (or L -formulas): (1) Let P_j be a predicate symbol of L and i_j be the corresponding natural number. If each of t_1, \dots, t_{i_j} is an L -term, then $P_j(t_1, \dots, t_{i_j})$ is an L -formula. This type of formula is called a **prime formula** (or **atomic formula**). (2) If A and B are L -formulas, then each of $\neg(A)$, $(A) \wedge (B)$, $(A) \vee (B)$, and $(A) \rightarrow (B)$ is an L -formula. (3) Let F be an L -formula and x_i be a bound variable that does not occur in F . Then an expression obtained by putting $(\)$ around F , replacing some occurrences in F of a free variable, say a_j , by x_i , and prefixing $\forall x_i$ or $\exists x_i$ is an L -formula. (4) The L -formulas are only those constructed by (1)–(3).

A formula that has no occurrence of a free variable is called a **closed formula**. The parentheses used in the formation of a formula may be omitted if no ambiguity arises thereby.

B. Structures

Let L be a specific language as described in the previous section. Then $\mathfrak{M} = [M : \rho; \sigma; \tau]$ defined by (1)–(4) below is called a **structure** for L (or L -structure).

(1) M is a nonempty set. (M is called the **universe** of \mathfrak{M} .)

(2) ρ is a mapping from L_4 into M .

(3) Let $L_5^i = \{f_j \mid \text{the number of arguments of } f_j \text{ is } i\}$. Then $L_5 = L_5^1 \cup L_5^2 \cup \dots \cup L_5^i \cup \dots$ provides a partition of L_5 . Let \mathfrak{F}_i be the set of all mappings from $M^i = M \times \dots \times M$ (i times) into M and σ_i be a mapping from L_5^i into \mathfrak{F}_i . We define σ for an arbitrary f of L_5 by $\sigma(f) = \sigma_i(f)$, where i is the number of arguments of f . Then σ is obviously a mapping from L_5 into $\bigcup_{i=1}^{\infty} \mathfrak{F}_i$.

(4) Decompose L_6 into $L_6^0 \cup L_6^1 \cup \dots \cup L_6^i \cup \dots$ as in (3). Let P_i be the set of all subsets of M^i and τ_i be a mapping from L_6^i into P_i , where P_0 is the set $\{M, \emptyset\}$ (\emptyset is the empty set). Then τ is defined for every i and for an arbitrary P of L_6^i by $\tau(P) = \tau_i(P)$.

If we denote $\rho(c)$ by \bar{c} , $\sigma(f)$ by \bar{f} , and $\tau(P)$ by \bar{P} , then we may understand that ρ is represented by $\bar{c}_0, \bar{c}_1, \dots$, σ is represented by $\bar{f}_0, \bar{f}_1, \dots$, and τ is represented by $\bar{P}_0, \bar{P}_1, \dots$. Therefore \mathfrak{M} is normally expressed as

$$\mathfrak{M} = [M : \bar{c}_0, \bar{c}_1, \dots; \bar{f}_0, \bar{f}_1, \dots; \bar{P}_0, \bar{P}_1, \dots].$$

C. Satisfiability

We fix not only a language L but also a structure \mathfrak{M} for L . Then the property that an L -formula is **satisfiable** is defined by the following procedure:

Let m, n, \dots stand for i -sequences of the elements of M , say $(m_0, m_1, \dots), (n_0, n_1, \dots), \dots$, called \mathfrak{M} -sequences. We write $m = n$ to indicate that each entry of m except the i th one is equal to the corresponding entry of n . Using these concepts, the value of an L -term at an \mathfrak{M} -sequence m , denoted by $t[m]$, is defined as follows:

(1) If t is a free variable a_j , then $t[m] = m_j$.

(2) If t is an individual constant c_j , then $t[m] = \bar{c}_j$.

(3) If t is of the form $f_j(t_1, \dots, t_i)$, then $t[m] = \bar{f}_j(t_1[m], \dots, t_i[m])$. If t is an L -term, then evidently $t[m]$ is an element of M .

Based on this definition of $t[m]$, the relation A is satisfiable by m in \mathfrak{M} , denoted by $\mathfrak{M}, m \models A$, is defined for an arbitrary L -formula A and an arbitrary \mathfrak{M} -sequence m as follows:

(1) $\mathfrak{M}, m \models P_j(t_1, \dots, t_i) \Leftrightarrow \langle t_1[m], \dots, t_i[m] \rangle \in \bar{P}_j$.

(2) $\mathfrak{M}, m \models \neg B \Leftrightarrow \mathfrak{M}, m \models B$ is false.

(3) $\mathfrak{M}, m \models B \wedge C \Leftrightarrow \mathfrak{M}, m \models B$ and $\mathfrak{M}, m \models C$.

(4) $\mathfrak{M}, m \models B \vee C \Leftrightarrow \mathfrak{M}, m \models B$ or $\mathfrak{M}, m \models C$.

(5) $\mathfrak{M}, m \models B \rightarrow C \Leftrightarrow \mathfrak{M}, m \models B$ implies $\mathfrak{M}, m \models C$.

(6) $\mathfrak{M}, m \models \forall x_j F(x_j) \Leftrightarrow \mathfrak{M}, n \models F(a_i)$ for an arbitrary n that satisfies $m = n$, where a_i has the least index among the free variables that do not occur in $F(x_j)$.

(7) $\mathfrak{M}, m \models \exists x_j F(x_j) \Leftrightarrow$ there exists an n such that $n = m$ and $\mathfrak{M}, n \models F(a_i)$, where a_i satisfies the same condition as in (6).

Following are some consequences of this definition.

(1) For an arbitrary L -formula A and an arbitrary \mathfrak{M} -sequence m , exactly one of $\mathfrak{M}, m \models A$ and $\mathfrak{M}, m \models \neg A$ holds.

(2) Let a_{j_1}, \dots, a_{j_i} include all free variables that occur in A , and let m and n be \mathfrak{M} -sequences for which $m_{j_1} = n_{j_1}, \dots, m_{j_i} = n_{j_i}$. Then $\mathfrak{M}, m \models A$ and $\mathfrak{M}, n \models A$ are equivalent.

So we may write $\mathfrak{M} \models A \left[\begin{smallmatrix} a_{j_1}, \dots, a_{j_i} \\ m_{j_1}, \dots, m_{j_i} \end{smallmatrix} \right]$ instead of $\mathfrak{M}, m \models A$, for any formula A whose free variables are among a_{j_1}, \dots, a_{j_i} , and any \mathfrak{M} -sequence m .

(3) If A is a closed formula, then for an arbitrary pair of sequences m and n , $\mathfrak{M}, m \models A$ and $\mathfrak{M}, n \models A$ are equivalent. Therefore, for a closed formula A , we may express the statement "for some (or, equivalently, for all) \mathfrak{M} -sequence m , $\mathfrak{M}, m \models A$ holds" by $\mathfrak{M} \models A$.

(4) Let a_i be an arbitrary variable that does not occur in $\forall x F(x)$ or $\exists x F(x)$. Then $\mathfrak{M}, m \models \forall x F(x)$ is equivalent to $\mathfrak{M}, n \models F(a_i)$ for an arbitrary n such that $n = m$. Likewise, $\mathfrak{M}, m \models \exists x F(x)$ is equivalent to the statement that there is an n such that $n = m$ and $\mathfrak{M}, n \models F(a_i)$.

D. Models

Here again we fix a language L . Let A be a closed L -formula and \mathfrak{M} an L -structure. If $\mathfrak{M} \models A$, then \mathfrak{M} is called a **model** of A .

Furthermore, if $\Gamma = \{A_0, A_1, \dots\}$ is an arbitrary set of closed formulas and $\mathfrak{M} \models A_i$ for all A_i in Γ , then the structure \mathfrak{M} is called a **model** of Γ .

(1) **Consistency.** Consider a logical system whose language is L . If there is a model of the set of all provable closed formulas of the system, then the system is i -consistent. In particular, the i -first-ordered predicate calculus is consistent.

(2) **Completeness.** A logical system is said to be **complete** if every closed formula that is satisfied in every structure is provable in the

system. In particular, the first-order predicate calculus is complete.

K. Gödel proved (2). Later L. Henkin gave an alternative proof whose essential idea contributed to proving the following proposition: If a set Γ of closed L -formulas is consistent, then there is a model of Γ . Henkin also introduced a (nonstandard) second-order semantics, relative to which the \dagger second-order predicate calculus is complete. This can be shown by extending Henkin's technique (for the first order) to the second-order language.

(3) Here we extend the language slightly by adding the second-order free predicate variables $\alpha_1^n, \alpha_2^n, \dots, \alpha_i^n, \dots$ ($n = 1, 2, \dots$) and the second-order bound predicate variables $\varphi_1^n, \varphi_2^n, \dots, \varphi_i^n, \dots$ ($n = 1, 2, \dots$), where n indicates the number of arguments of a variable. Otherwise the definition of the language is the same as for the case of the first-order predicate calculus. For simplicity, however, we assume that there are no individual constants, function symbols, or predicate symbols.

The structure is defined as follows: Put $\mathfrak{M} = [M: S_1, S_2, \dots, S_n, \dots]$, where M is a nonempty set and S_n is a set of subsets of $M \times \dots \times M$ (n times). An \mathfrak{M} -sequence m is defined as before, and s_n denotes $(s_1^n, s_2^n, \dots, s_i^n, \dots)$, where each s_i^n is a member of S_n . The concept of satisfiability is defined as follows:

$\mathfrak{M}, (m, s_1, \dots, s_n, \dots) \models \alpha_j^n(x_{i_1}, \dots, x_{i_n}) \Leftrightarrow (m_{i_1}, \dots, m_{i_n}) \in s_j^n$.

$\mathfrak{M}, (m, s_1, \dots, s_n, \dots) \models \forall \varphi_k^n A(\varphi_k^n) \Leftrightarrow$ for an arbitrary s'_n for which $s'_n \neq s_n$, $\mathfrak{M}, (m, s_1, \dots, s'_n, \dots) \models A(\alpha_j^n)$, where α_j^n has the smallest index among the free predicate variables that do not occur in $A(\varphi_k^n)$.

$\mathfrak{M}, (m, s_1, \dots, s_n, \dots) \models \exists \varphi_k^n A(\varphi_k^n) \Leftrightarrow$ there exists an s'_n such that $s'_n \neq s_n$ and $\mathfrak{M}, (m, s_1, \dots, s'_n, \dots) \models A(\alpha_j^n)$, where α_j^n satisfies the same condition as in the previous clause.

Satisfiability for other cases is defined as for first-order predicate language. A structure \mathfrak{M} is called **normal** if all axioms of the second-order predicate calculus are true in \mathfrak{M} .

Completeness of the second-order predicate calculus: Every closed formula that is satisfiable in all normal structures is provable in the second-order predicate calculus.

(4) Let the cardinality of L_4 be τ , and Γ be an arbitrary set of closed L -formulas. If Γ has a model, then Γ has a model of cardinality $\max(\tau, \aleph_0)$. This follows from Henkin's method. Historically, however, it was first proved by Th. Skolem and L. Löwenheim for a special case, and was later generalized by A. I. Mal'tsev and A. Robinson.

(5) The following results are all due to A. Tarski and R. L. Vaught.

Definition 1. Two L -structures \mathfrak{M} and \mathfrak{N} are said to be **elementarily (arithmetically) equivalent** if for an arbitrary closed L -formula A , $\mathfrak{M} \models A \Leftrightarrow \mathfrak{N} \models A$.

Definition 2. Let

$\mathfrak{M} = [M: q_0, q_1, \dots; g_0, g_1, \dots; Q_0, Q_1, \dots]$

and

$\mathfrak{N} = [N: r_0, r_1, \dots; h_0, h_1, \dots; R_0, R_1, \dots]$

be two structures. \mathfrak{M} is an **elementary extension** of \mathfrak{N} if the following two conditions are satisfied: (i) $M \supset N$; $q_j = r_j$ ($j = 0, 1, \dots$); the restriction of g_j to N is identical to h_j ($j = 0, 1, \dots$); the restriction of Q_j to N is identical to R_j ($j = 0, 1, \dots$). (If this condition holds, then \mathfrak{M} is said to be an extension of \mathfrak{N} .) (ii) For an arbitrary L -formula A and an arbitrary \mathfrak{N} -sequence n , if $\mathfrak{N}, n \models A$ then $\mathfrak{M}, n \models A$.

Theorem 1. Let \mathfrak{M} be an extension of \mathfrak{N} . A necessary and sufficient condition for \mathfrak{M} to be an elementary extension of \mathfrak{N} is that for an arbitrary L -formula of the form $\exists x F(x)$ and an arbitrary \mathfrak{N} -sequence n , if $\mathfrak{M}, n \models \exists x F(x)$, then there is some element n of N such that for the \mathfrak{M} -sequence m for which $m \stackrel{i}{=} n$ and $m_i = n$, $\mathfrak{M}, m \models F(a_i)$, where a_i is an arbitrary free variable that does not occur in $F(x)$.

Theorem 2. Here we place a condition on L that each set of symbols be at most countable and arranged in the ω -type (\rightarrow 312 Ordinal Numbers). Let the cardinality of the universe M of \mathfrak{M} be an infinite cardinal α , M' be a subset of M of cardinality ϵ , and b be an infinite cardinal that satisfies $\epsilon \leq b \leq \alpha$. Then there exists an L -structure \mathfrak{N} whose universe N has cardinality b and such that $M' \subset N$ and \mathfrak{M} is an elementary extension of \mathfrak{N} .

Theorem 3. Suppose that L satisfies the same condition as in Theorem 2. Let the cardinality of the universe M of \mathfrak{M} be α (α is an infinite cardinal) and b be a cardinal for which $\alpha \leq b$. Then there exists an L -structure \mathfrak{N} that is a proper elementary extension of \mathfrak{M} and whose universe has cardinality b .

E. Ultraproducts

Assume that for a set of L -structures Σ and a set of indices I , there is a mapping θ from I onto Σ . If α is a member of I , \mathfrak{M} is a member of Σ , and $\theta(\alpha) = \mathfrak{M}$, then \mathfrak{M} may be denoted by \mathfrak{M}^α . It should be noted that there may be more than one α corresponding to the same structure. If D is a \dagger maximal filter of I and \mathfrak{M}^α

is expressed as

$$\mathfrak{M}^\alpha = [M^\alpha, \bar{c}^\alpha, \dots, \bar{f}^\alpha, \dots, \bar{P}^\alpha, \dots],$$

then $\prod_{\alpha \in I} M^\alpha$ is defined by

$$\prod_{\alpha \in I} M^\alpha = \{ \varphi \mid \varphi \text{ is a mapping from } I \text{ into } \bigcup_{\alpha \in I} M^\alpha, \text{ where } \varphi(\alpha) \in M^\alpha \}.$$

For any two elements φ and ψ of $\prod_{\alpha \in I} M^\alpha$, $\varphi \stackrel{D}{=} \psi$ is defined by

$$\varphi \stackrel{D}{=} \psi \Leftrightarrow \{ \alpha \mid \varphi(\alpha) = \psi(\alpha) \} \in D.$$

Then $\varphi \stackrel{D}{=} \psi$ is an equivalence relation between the elements of $\prod_{\alpha \in I} M^\alpha$. Furthermore, the set $\prod_{\alpha \in I} M^\alpha$ partitioned by $\stackrel{D}{=}$ is expressed by $\prod_{\alpha \in I} M^\alpha / D$, and each element m of $\prod_{\alpha \in I} M^\alpha / D$ is expressed by $m = [\varphi]$, where φ is a representing element of m .

Next we define an operator that produces a new structure from Σ . Put $M = \prod_{\alpha \in I} M^\alpha / D$. For an individual constant c of L , let $\bar{c} = [\varphi]$, where $\varphi(\alpha) = c^\alpha$ for every α . For an n -ary function f of L and arbitrary elements $m_1 = [\varphi_1], \dots, m_n = [\varphi_n]$ of M , define $\bar{f}(m_1, \dots, m_n) = [\psi]$, where $\psi(\alpha) = \bar{f}^\alpha(\varphi_1(\alpha), \dots, \varphi_n(\alpha))$ for every α . For an n -ary predicate P of L , define $\langle m_1, \dots, m_n \rangle \in \bar{P}$ by

$$\langle m_1, \dots, m_n \rangle \in \bar{P} \Leftrightarrow \{ \alpha \mid \langle \varphi_1(\alpha), \dots, \varphi_n(\alpha) \rangle \in P^\alpha \} \in D.$$

According to these definitions, put

$$\mathfrak{M} = [M, \bar{c}, \dots, \bar{f}, \dots, \bar{P}, \dots],$$

and denote it by $\prod_{\alpha \in I} \mathfrak{M}^\alpha / D$, called the **ultraproduct** of $\{\mathfrak{M}^\alpha\}_{\alpha \in I}$ (with respect to D). \mathfrak{M} is an L -structure.

Fundamental Theorem of Ultraproducts.

Let $\mathfrak{M} = \prod_{\alpha \in I} \mathfrak{M}^\alpha / D$ be the ultraproduct of $\{\mathfrak{M}^\alpha\}_{\alpha \in I}$, $m = (m_1, m_2, \dots)$ be an \mathfrak{M} -sequence, φ_i be a representing element of m_i , and A be an arbitrary formula. Then $\mathfrak{M}, m \models A \Leftrightarrow \{ \alpha \mid \mathfrak{M}^\alpha, (\varphi_1(\alpha), \varphi_2(\alpha), \dots) \models A \} \in D$.

By using this fundamental theorem we have the following result: Suppose that Γ is a set of closed formulas in L such that every finite subset of Γ has a model. Let I be the set of all the finite subsets of Γ . For each $\alpha \in I$ and each $A \in \Gamma$, let \mathfrak{M}^α be a model of α and \hat{A} the set of all the finite sets in I which contain A as a member. Let $F = \{ \hat{A} \mid A \in \Gamma \}$. Since F has the finite intersection property, there is a maximal filter D such that $D \supseteq F$. Let \mathfrak{M} be the ultraproduct $\prod_{\alpha \in I} \mathfrak{M}^\alpha / D$. Since \hat{A} is a subset of the set $\{ \alpha \mid \mathfrak{M}^\alpha \models A \}$ and \hat{A} belongs to D , the set $\{ \alpha \mid \mathfrak{M}^\alpha \models A \}$ belongs to D , for each $A \in \Gamma$. Hence we have that M is a model of Γ , by the foregoing fundamental theorem of ultra-

products. Therefore, we have the following theorem.

Compactness Theorem. A set Γ of closed formulas has a model if and only if every finite subset of it has a model.

In the case where all the structures \mathfrak{M}^α coincide with the single structure \mathfrak{R} , the ultraproduct of $\{\mathfrak{M}^\alpha\}_{\alpha \in I}$ (with respect to D) may be written \mathfrak{R}^I / D and called the **ultrapower** of \mathfrak{R} (with respect to D).

Let

$$\mathfrak{M} = [M; q_0, q_1, \dots; g_0, g_1, \dots; Q_0, Q_1, \dots]$$

and

$$\mathfrak{N} = [N; r_0, r_1, \dots; h_0, h_1, \dots; R_0, R_1, \dots]$$

be two structures. \mathfrak{M} and \mathfrak{N} are said to be **isomorphic** if there is a bijection f from M to N such that the following three conditions hold: (i) $f(q_0) = r_0, f(q_1) = r_1, \dots$ (ii) The sequences g_0, g_1, \dots and h_0, h_1, \dots are of the same type and $f(g_j(a_1, \dots, a_n)) = h_j(f(a_1), \dots, f(a_n))$ holds for every n -tuple a_1, \dots, a_n in M . (iii) The sequences Q_0, Q_1, \dots and R_0, R_1, \dots are of the same type and $R_j = \{ \langle f(a_1), \dots, f(a_n) \rangle \mid \langle a_1, \dots, a_n \rangle \in Q_j \}$.

Let j be the function from N to N^I / D defined by $j(a) = [\varphi_a]$ for each $a \in N$, where φ_a is the constant function from I to N such that $\varphi_a(\alpha) = a$ for each $\alpha \in I$. Let \mathfrak{M} be the substructure of \mathfrak{N}^I / D whose universe is the range of j . Then j is an isomorphism of \mathfrak{N} to \mathfrak{M} . In the following we identify a and $j(a)$ for each $a \in N$. Then \mathfrak{N} is an elementary substructure of \mathfrak{N}^I / D by the fundamental theorem of ultraproducts.

If \mathfrak{M} and \mathfrak{N} are isomorphic, then \mathfrak{M} and \mathfrak{N} are elementarily equivalent. By using this fact and the fundamental theorem of ultraproducts, we have the following result. Let \mathfrak{M} and \mathfrak{N} be two structures. If there is a nonempty set I and a maximal filter D on I such that \mathfrak{M}^I / D and \mathfrak{N}^I / D are isomorphic, then \mathfrak{M} and \mathfrak{N} are elementarily equivalent. H. J. Keisler proved the converse of this proposition by using the G.C.H. (generalized continuum hypothesis), and later S. Shelah proved it without the G.C.H. **Keisler-Shelah isomorphism theorem:** Let \mathfrak{M} and \mathfrak{N} be two structures. Then \mathfrak{M} and \mathfrak{N} are elementarily equivalent if and only if there is a nonempty set I and a maximal filter D on I such that \mathfrak{M}^I / D and \mathfrak{N}^I / D are isomorphic.

The ultraproduct operation has various applications in number theory, algebraic geometry, and analysis. Here we give an example due to J. Ax and S. Kochen. Let P be the set of prime numbers. Let \mathbb{Q}_p and $\mathbb{Z}_p((t))$ be the field of p -adic numbers and the field of formal

power series over $Z_p = \{0, 1, \dots, p-1\}$ for each p in P , respectively. **Ax-Kochen isomorphism theorem:** Suppose that D is a nonprincipal maximal filter on P . Then $\prod_{p \in P} \mathbf{Q}_p/D$ and $\prod_{p \in P} Z_p((t))/D$ are isomorphic.

As an immediate consequence of this theorem, we have the following partial solution of Artin's conjecture on Diophantine equations.

Theorem: For each positive integer d , there exists a finite set Y of primes such that every homogeneous polynomial $f(t_1, \dots, t_n)$ of degree d over \mathbf{Q}_p , with $n > d^2$, has a nontrivial zero in \mathbf{Q}_p for every $p \notin Y$ (\rightarrow 118 Diophantine Equations).

We give another example in nonstandard analysis. A. Robinson developed the general theory of nonstandard analysis in [10]. Here we explain a theorem due to A. R. Bernstein. Let X be a nonempty set and $U(X)$ be the smallest transitive set (i.e., $a \in b$ and $b \in U(X)$ imply $a \in U(X)$) which has X as a member and is closed under the following operations: pairing, union, power set, and subset operation (i.e., $a \in U(X)$ and $b \subset a$ imply $b \in U(X)$). Let L be the first-order predicate logic with equality whose set of nonlogical constants consists of a binary predicate symbol \in and individual constant symbols c_a for $a \in U(X)$ (\rightarrow 411 Symbolic Logic F). Then the first-order structure $\mathfrak{R} = [U(X); a(a \in U(X)); E]$ is an L -structure, where E is the \in relation on the set $U(X)$. Let $\mathfrak{M} = [U(X)^I/D; a(a \in U(X)); E^I/D]$ be the ultrapower \mathfrak{R}^I/D of \mathfrak{R} with respect to a nonprincipal ultrafilter D on a set I . For each $a \in U(X)$, let a^* be the set of all elements $[\varphi]$ in $U(X)^I/D$ such that $\{i \in I \mid \varphi(i) \in a\} \in D$. Then a is a proper subset of a^* if a is infinite. Since \mathfrak{R} and \mathfrak{M} are elementarily equivalent, these two sets a and a^* have common first-order properties in the following sense: for each formula $\Phi(a_0)$ in L ,

$$(\forall b \in a)(\mathfrak{R} \models \Phi[b^a]) \Leftrightarrow (\forall b \in a^*)(\mathfrak{M} \models \Phi[b^{a_0}]).$$

From this it follows that if r is a relation on a set a in \mathfrak{R} , then r^* is a relation on the set a^* ; and if f is a mapping from a to b in \mathfrak{R} , then f^* is a mapping from a^* to b^* . Hence a^* is a mathematical object which greatly resembles a . By using this type of resemblance between a and a^* we have the following result.

Let H be a Hilbert space over the complex number field \mathbf{C} such that $\dim(H) = \omega$ and let T be a bounded linear operator on H . Let $X = H \cup \mathbf{C}$ and consider the first-order structure \mathfrak{M} as above. Since \mathbf{R} (the set of all real numbers) and \mathbf{N} (the set of all natural numbers) are infinite sets which belong to $U(X)$, \mathbf{R}^* and \mathbf{N}^* have elements which do not belong to \mathbf{R} and \mathbf{N} , respectively. Such elements are called **nonstandard real numbers** and **nonstandard natural numbers**, respectively. By the fundamental

theorem of ultraproducts we can conclude that there are many nonstandard real numbers α such that $0 < {}^*\alpha < {}^*a$ in \mathbf{R}^* for any $a \in \mathbf{R}$. Such a nonstandard real number α is called an **infinitesimal real number**. Since the norm operator $\| \cdot \|$ is a mapping from H to \mathbf{R} , $\| \cdot \|$ is a mapping from H^* to \mathbf{R}^* . If $\|x\|$ ($x \in H^*$) is infinitesimal, then x is said to be **infinitesimal** in H^* . Let S be the set of all linear subspaces of H . For a linear subspace K of H^* that is contained in S^* let K° be the set of elements $x \in H$ such that $x - x_0$ is infinitesimal for some x_0 in K . Then K° is a closed linear subspace of H . Let $e = \{e_i\}_{i \in \mathbf{N}}$ be an orthonormal basis of the Hilbert space H ; e can be considered as a mapping from \mathbf{N} to H , and hence $e^* = \{e_j\}_{j \in \mathbf{N}^*}$ is a mapping from \mathbf{N}^* to H^* . For each $j \in \mathbf{N}^*$, let H_j be the linear subspace of H^* spanned by $\{e_k \mid k \leq j\}$. For a given bounded linear operator T on H , T^* is a linear operator on H^* . We define $T_j = P_j T^* P_j$, where P_j is the projection from H^* to H_j . Since $\dim(H_j) = j$ is a (nonstandard) natural number, there exists a tower $J_{0j} \subset J_{1j} \subset \dots \subset J_{jj} = H_j$ of closed, T_j -invariant linear subspaces of H_j such that $\dim(J_{kj}) = k$ ($k \leq j$). Then J_{kj}° is a closed, T -invariant linear subspace of H . If there is a polynomial $p(x)$ such that $p(T)$ is a compact operator, then we get a nonstandard natural number j such that J_{kj}° is a proper subspace of H for some $k \leq j$. This gives the following result, which is an affirmative solution of a problem of K. Smith and P. R. Halmos.

Theorem (Bernstein [4]): Let T be a bounded linear operator on an infinite-dimensional Hilbert space H over the complex numbers and let $p(x) \neq 0$ be a polynomial with complex coefficients such that $p(T)$ is compact. Then T leaves invariant at least one closed linear subspace of H other than H or $\{0\}$.

F. Categoricity in Powers

Let Γ be a set of closed formulas in a first-order language L which has a designated binary predicate symbol P_0 . In the following, we assume that the interpretation \bar{P}_0 of P_0 by \mathfrak{M} is the equality relation on the universe of \mathfrak{M} for every L -structure \mathfrak{M} . Γ is said to be **categorical** if all the models of Γ are isomorphic. By Theorem 3 in Section D, any Γ having a model of infinite cardinality is not categorical. Hence, there exists no interesting Γ which is categorical. Therefore, we consider the weaker notion of **categoricity in powers**. Let κ be an infinite cardinal and $n(\Gamma, \kappa)$ be the number of nonisomorphic models of Γ of cardinality κ . Then Γ is said to be categorical in κ if $n(\Gamma, \kappa) = 1$, i.e., if all the models of Γ of cardinality κ are isomorphic. There exist many

interesting Γ 's which are categorical in κ for some κ . For example, the set of axioms of algebraically closed fields of characteristic 0 is categorical in \aleph_1 , and the set of axioms of dense linear orderings without endpoints is categorical in \aleph_0 . With respect to this notion, J. Łoś conjectured that if Γ is categorical in κ for some $\kappa > \bar{L}$ (the cardinality of L), then Γ is categorical in κ for all $\kappa > \bar{L}$. This conjecture was solved affirmatively by M. Morley in the case $\bar{L} = \aleph_0$, and later by S. Shelah in the general case. **Theorem (1):** Let Γ be a set of closed formulas in L . Then Γ is categorical in κ for some $\kappa > \bar{L}$ if and only if Γ is categorical in κ for all $\kappa > L$. We also have the following interesting theorem, due to J. T. Baldwin and A. H. Lachlan. **Theorem (2):** Let Γ be a set of closed formulas in L such that $L = \aleph_0$. If Γ is categorical in \aleph_1 , then $n(\Gamma, \aleph_0) = 1$ or \aleph_0 .

As for $n(\Gamma, \kappa)$ there are two famous conjectures due to R. Vaught and others. A set Γ of closed formulas in L is called **complete** in L if it has a model and, for any closed formula A in L , either $A \in \Gamma$ or $\neg A \in \Gamma$.

Conjecture 1. There is no complete set Γ of closed formulas in L such that $\aleph_0 < n(\Gamma, \aleph_0) < 2^{\aleph_0}$.

Conjecture 2. There is no finite set Γ of closed formulas in L such that $n(\Gamma, \aleph_0) = \aleph_0$ and $n(\Gamma, \aleph_1) = 1$.

G. Omitting Type Theorem

By a_0 -formulas (or a -formulas), we mean formulas which have no free variables except a_0 . Suppose that \mathfrak{M} is an L -structure and Σ is a set of a -formulas in L . Then, \mathfrak{M} realizes Σ if there is an element m of the universe of \mathfrak{M} such that $\mathfrak{M} \models A[m]$ for all A in Σ , and \mathfrak{M} omits Σ if \mathfrak{M} does not realize Σ . For example, if L is a first-order language such that $L_4 = \{c_0, c_1\}$, $L_5 = \{f_0, f_1\}$, $L_6 = \{P_0, P_1\}$, where f_0, f_1, P_0, P_1 are all binary, and \mathfrak{M} is the L -structure $[\mathfrak{M}; 0, 1; +, x; =, <]$, where \mathfrak{M} is the set of natural numbers, and $\Sigma = \{P_1(\bar{n}, a) \mid n=0, 1, 2, \dots\}$, where $\bar{0} = c_0$, $\bar{1} = f_0(\bar{0}, c_1)$, \dots , $\bar{n+1} = f_0(\bar{n}, c_1)$, \dots , then clearly \mathfrak{M} omits Σ . Also, if \mathfrak{M} is an L -structure, m is an element of the universe of \mathfrak{M} , and $\Sigma_m = \{A \mid A \text{ is an } a\text{-formula such that } \mathfrak{M} \models A[m]\}$, then clearly \mathfrak{M} realizes Σ_m , where Σ_m is called the type of m in \mathfrak{M} . Suppose that Γ is a set of closed formulas in L . Then, by the completeness theorem of L , we can easily see that Γ has a model realizing Σ if and only if $\Gamma \cup \Sigma$ is consistent. On the other hand, it is rather difficult to obtain a necessary and sufficient condition for Γ to have a model omitting Σ . The following is a sufficient con-

dition: Γ is said to locally omit Σ if there is no a -formula A such that $\Gamma \cup \{A\}$ is consistent and, for any formula B in Σ , $\Gamma \cup \{A, \neg B\}$ is not consistent.

Theorem. Suppose that Γ is a consistent set of closed formulas in a countable language L , and Σ is a set of a_0 -formulas in L . If Γ locally omits Σ , then Γ has a countable model which omits Σ . Also, if Γ has a model of power greater than \beth_ξ omitting Σ for each $\xi < \omega_1$, then Γ has a model omitting Σ in each infinite power, where \beth_ξ is defined by $\beth_0 = \aleph_0$, $\beth_{\xi+1} = 2^{\beth_\xi}$, $\beth_\sigma = \sup_{\xi < \sigma} \beth_\xi$ if σ is a limit ordinal.

References

- [1] J. Ax and S. Kochen, Diophantine problems over local fields I, *Amer. J. Math.*, 87 (1965), 605–630.
- [2] J. Ax and S. Kochen, Diophantine problems over local fields II, *Amer. J. Math.*, 87 (1965), 630–648.
- [3] J. Ax and S. Kochen, Diophantine problems over local fields III, *Ann. Math.*, (2) 83 (1966), 437–456.
- [4] A. R. Bernstein, Non-standard analysis, *Studies in Model Theory* 8, M. D. Morley (ed.), Mathematical Association of America, 1973, 35–58.
- [5] C. C. Chang and H. J. Keisler, *Model theory*, North-Holland, 1973.
- [6] S. C. Kleene, *Introduction to metamathematics*, Van Nostrand, 1952.
- [7] M. Machover and J. Hirschfeld, *Lectures on non-standard analysis*, Lecture notes in math. 94, Springer, 1969.
- [8] M. Morley, Categoricity in powers, *Trans. Am. Math. Soc.*, 114 (1965), 514–538.
- [9] M. Morley, Omitting types of elements, *The Theory of Models*, J. W. Addison et al. (eds.), North-Holland, 1965, 265–273.
- [10] A. Robinson, *Non-standard analysis*, North-Holland, 1966.
- [11] G. E. Sacks, *Saturated model theory*, Benjamin, 1972.
- [12] S. Shelah, *Classification theory*, North-Holland, 1978.

277 (III.23) Modules

A. General Remarks

In this article, we consider mainly modules with operator domain (Section C), in particular modules over a \dagger ring. Modules over a

field are linear spaces (\rightarrow 256 Linear Spaces). Modules over a commutative ring are important in algebraic geometry (\rightarrow 16 Algebraic Varieties, 67 Commutative Rings, 284 Noetherian Rings). The theory of modules over a \dagger group ring can be identified with the theory of linear representations of a group (\rightarrow 362 Representations). Modules without operator domain may be regarded as modules over the ring \mathbf{Z} of rational integers, and the theory of finitely generated Abelian groups can be generalized to the theory of modules over a \dagger principal ideal domain (\rightarrow 52 Categories and Functors, 200 Homological Algebra).

B. Modules

A **module** (without operator domain) is a \dagger commutative group M whose law of composition is written additively: $a + b = b + a$ ($a, b \in M$); the \dagger identity element is denoted by 0, and the inverse element of a by $-a$. Every subgroup H of M is a normal subgroup. For any $a \in M$, the left and right cosets of H containing the element a are identical: $H + a = a + H$ (\rightarrow 190 Groups A).

In the set N^M of all mappings of a set M to a module N , we define an addition by the sums of values: $(f + g)(x) = f(x) + g(x)$. Then N^M forms a module. The set $\text{Hom}(M, N)$ of all homomorphisms of a module M to a module N forms a subgroup of the module N^M , called the **module of homomorphisms** of M to N . The composite of homomorphisms is a homomorphism. Hence the set $\text{Hom}(M, M) = \mathcal{E}(M)$ of all endomorphisms of M forms a \dagger ring with respect to the addition and the multiplication defined by composition; this is called the **endomorphism ring** of M . The \dagger unity element of $\mathcal{E}(M)$ is the identity mapping of M , and the \dagger invertible elements of $\mathcal{E}(M)$ are the automorphisms of M .

Let $\{x_\lambda\}_{\lambda \in \Lambda}$ be a family of elements in a module M . The sum $\sum_{\lambda \in \Lambda} x_\lambda$ is well defined if $x_\lambda = 0$ ($\lambda \in \Lambda$) except for a finite number of λ . For any family $\{N_\lambda\}_{\lambda \in \Lambda}$ of subsets of M , $\sum_{\lambda \in \Lambda} N_\lambda$ denotes the set of all elements of the form $\sum_{\lambda \in \Lambda} x_\lambda$ ($x_\lambda \in N_\lambda$), where $x_\lambda = 0$ except for a finite number of λ . If all the N_λ are subgroups of M , then $N = \sum_{\lambda \in \Lambda} N_\lambda$ is also a subgroup, called the **sum** of $\{N_\lambda\}_{\lambda \in \Lambda}$. If every element of N can be written uniquely in the form $\sum_{\lambda \in \Lambda} x_\lambda$ ($x_\lambda \in N_\lambda$), N is called the **direct sum** of $\{N_\lambda\}_{\lambda \in \Lambda}$. When the N_λ are subgroups, this is equivalent to the condition that $N_\lambda \cap \sum_{\lambda \neq \mu \in \Lambda} N_\mu = \{0\}$ ($\lambda \in \Lambda$).

C. Modules with Operator Domain

Suppose that we are given a set A and a module M . If with each pair of elements $a \in A$

and $x \in M$ there is associated a unique element $ax \in M$ satisfying the condition (1) $a(x + y) = ax + ay$ ($a \in A$; $x, y \in M$), we say that A is an **operator domain** of M and M is a **module with operator domain A** (**module over A** or **A -module**) (\rightarrow 190 Groups E). The mapping $A \times M \rightarrow M$ given by $(a, x) \rightarrow ax$ is called the **\dagger operation** of A on M . Any $a \in A$ induces an endomorphism $a_M: x \rightarrow ax$ of M as a module (not as an A -module). To give the structure of an A -module to a module M amounts to giving a mapping $A \rightarrow \mathcal{E}(M)$ ($a \rightarrow a_M$).

If N is a subgroup of an A -module M such that $ax \in N$ for any $a \in A$ and $x \in N$, then N forms an A -module, called an **A -submodule** (or **allowed submodule**) of M . If $\{N_\lambda\}_{\lambda \in \Lambda}$ is a family of A -submodules of an A -module M , then the intersection $\bigcap_{\lambda \in \Lambda} N_\lambda$ and the sum $\sum_{\lambda \in \Lambda} N_\lambda$ are both A -submodules of M .

Let R be an \dagger equivalence relation in an A -module M such that if $a \in A$ and xRy , then $axRay$. Then R is said to be **compatible with the operation** of A . In this case, an operation of A is induced on the quotient set M/R . Moreover, if R is compatible with the addition, namely, xRx' and yRy' imply $(x + y)R(x' + y')$, then M/R forms an A -module, called a **factor A -module** of M . The equivalence class N containing 0 is an A -submodule of M , and M/R coincides with M/N .

D. Modules over a Group or a Ring

If a (multiplicative) group structure is given to the operator domain A of a module M , we always assume (in addition to condition (1) in Section C) that the following two conditions are satisfied: (2) $(ab)x = a(bx)$; (3) $1x = x$ ($a, b \in A$, $x \in M$).

If a ring structure is given to A , we always assume (besides conditions (1) and (2)) that the following condition holds: (4) $(a + b)x = ax + bx$ ($a, b \in A$, $x \in M$). This means that the mapping $A \rightarrow \mathcal{E}(M)$ ($a \rightarrow a_M$) is a \dagger ring homomorphism. If the ring A has unity element 1, and $1_M = \text{identity mapping}$ (namely, condition (3) holds), then the A -module M is called **unitary**. We consider only unitary A -modules. Any module M can be regarded as a \mathbf{Z} -module or as an $\mathcal{E}(M)$ -module.

When M is a module over a ring A , an element of A is called a **scalar**, A itself is called the **ring of scalars** (**basic ring** or **ground ring**), and the operation $A \times M \rightarrow M$ is called the **scalar multiplication**. The elements ax ($a \in A$) are called **scalar multiples** of x , and the totality of these elements is denoted by Ax . Let $\{x_\lambda\}_{\lambda \in \Lambda}$ be a family of elements in M . An element of the form $\sum_{\lambda \in \Lambda} a_\lambda x_\lambda$, where the a_λ are elements of A and equal to 0 except for a finite number

of λ , is called a linear combination of $\{x_\lambda\}_{\lambda \in \Lambda}$. The set N of linear combinations of $\{x_\lambda\}_{\lambda \in \Lambda}$ is the smallest A -submodule of M containing all the x_λ ($\lambda \in \Lambda$) and is equal to the sum $\sum_{\lambda \in \Lambda} Ax_\lambda$. The A -module N is said to be **generated** by $\{x_\lambda\}_{\lambda \in \Lambda}$, and $\{x_\lambda\}_{\lambda \in \Lambda}$ is called a **system of generators** of N . A module having a finite number of generators is said to be **finitely generated** (of **finite type** or simply **finite**). The module Ax generated by a single element x is called **monomial**. If A is a * field (which may be noncommutative), an A -module is a linear space over A (\rightarrow 256 Linear Spaces).

Let a be an element of an A -module M . If there exists a nonzero divisor λ of A such that $\lambda a = 0$, then a is called a **torsion element** of M . We say M is a **torsion A -module** if every element of M is a torsion element, and M is **torsion free** if M has no torsion element other than 0. An element a of M is called **divisible** if for any nonzero divisor $\lambda \in A$ there exists an element $b \in M$ such that $a = \lambda b$. M is called a **divisible A -module** if every element of M is divisible.

Strictly speaking, the A -modules we have considered so far are called **left A -modules**. If we define the operation of A on M by xa ($a \in A, x \in M$) instead of ax and modify conditions (1)–(4) appropriately (in particular, condition (2) becomes $x(ab) = (xa)b$), then M is called a **right A -module**. If A° is a group or ring anti-isomorphic to A , then a left A -module can be naturally identified with a right A° -module. If A is a commutative group or ring, we can disregard the distinction between left and right A -modules.

Let A and B be groups or rings. Sometimes we consider an A -module structure and a B -module structure simultaneously on the same module M . If the operations of A and B commute with each other, namely, $a(bx) = b(ax)$ ($a \in A, b \in B, x \in M$), it is convenient to put one of the operations to the right. If M has a left A -module structure and a right B -module structure, satisfying condition (5) $(ax)b = a(xb)$, then M is called an **A - B -bimodule**. If G is a group and K is a commutative ring, the G - K -bimodule structure is equivalent to the left $K[G]$ -module structure, where $K[G]$ is the * group ring.

E. Operator Homomorphisms

A homomorphism $f: M \rightarrow N$ of A -modules M and N such that $f(ax) = af(x)$ ($a \in A, x \in M$) is called an **A -homomorphism (operator homomorphism or allowed homomorphism)**. If A is a ring, f is also called an **A -linear mapping**. Regarding A as an A -module, an A -linear mapping $M \rightarrow A$ is called a **linear form** on M .

The composite of A -homomorphisms is an A -homomorphism.

Let $f: M \rightarrow L$ be an A -homomorphism of A -modules. The A -submodule $\text{Im } f = f(M)$ of L is called the **image** of f , and the A -submodule $\text{Ker } f = \{x \in M, f(x) = 0\}$ of M is called the **kernel** of f . $\text{Coim } f = M/\text{Ker } f$ is called the **co-image** of f , and $\text{Coker } f = L/\text{Im } f$ the **cokernel** of f . The binary relation xRy on M defined by $f(x) = f(y)$ ($x, y \in M$) coincides with the equivalence relation defined by $x - y \in N = \text{Ker } f$ ($x, y \in M$), and the mapping f induces an A -isomorphism $\bar{f}: M/N \rightarrow f(M)$.

A sequence of A -homomorphisms of A -modules M_n ($n \in \mathbb{Z}$)

$$\dots \rightarrow M_{n-1} \xrightarrow{f_{n-1}} M_n \xrightarrow{f_n} M_{n+1} \rightarrow \dots$$

is called an **exact sequence** if $\text{Im } f_{n-1} = \text{Ker } f_n$ for all n . The A -module $\{0\}$ is denoted by 0. Exactness of $0 \rightarrow N \xrightarrow{f} M$ or $M \xrightarrow{g} L \rightarrow 0$ means that the mapping $f: N \rightarrow M$ is injective or the mapping $g: M \rightarrow L$ is surjective, respectively. In an * inductive (* projective) system $\{M_\lambda, f_{\lambda\mu}\}$ of A -modules, where every $f_{\lambda\mu}: M_\lambda \rightarrow M_\mu$ ($\lambda < \mu$) ($\lambda > \mu$) is an A -homomorphism, the limit $M = \varinjlim M_\lambda$ ($\varprojlim M_\lambda$) is also an A -module. If $0 \rightarrow L_\lambda \rightarrow M_\lambda \rightarrow N_\lambda \rightarrow 0$ is exact for every λ and

$$\begin{array}{ccccc} L_\lambda & \rightarrow & M_\lambda & \rightarrow & N_\lambda \\ \downarrow & & \downarrow & & \downarrow \\ L_\mu & \rightarrow & M_\mu & \rightarrow & N_\mu \end{array}$$

is a * commutative diagram, then $0 \rightarrow \varinjlim L_\lambda \rightarrow \varinjlim M_\lambda \rightarrow \varinjlim N_\lambda \rightarrow 0$ is also exact. For the projective limit, however, we can only state the exactness of $0 \rightarrow \varprojlim L_\lambda \rightarrow \varprojlim M_\lambda \rightarrow \varprojlim N_\lambda$.

The set of all A -homomorphisms of an A -module M to an A -module N , denoted by $\text{Hom}_A(M, N)$, is a subgroup of the module $\text{Hom}(M, N)$, and is called the **module of A -homomorphisms**. The set $\text{Hom}_A(M, M) = \mathcal{E}_A(M)$ of all A -endomorphisms of an A -module M forms a * subring of the ring $\mathcal{E}(M)$ and coincides with the set of all elements commuting with any a_M ($a \in A$). We denote by $GL(M)$ the group of all * invertible elements in $\mathcal{E}_A(M)$. If A is a commutative ring, $\text{Hom}_A(M, N)$ can be regarded as an A -module by defining $(af)(x) = af(x)$, namely, $af = a_M \circ f$. In particular, $\mathcal{E}_A(M)$ is an * associative algebra over A . If M is an A - B -bimodule, $\text{Hom}_A(M, N)$ forms a left B -module by $(bf)(x) = f(xb)$. If N is an A - B -bimodule, $\text{Hom}_A(M, N)$ forms a right B -module by $(fb)(x) = f(x)b$.

F. Direct Products and Direct Sums

In the Cartesian product $P = \prod_{\lambda \in \Lambda} M_\lambda$ of a family $\{M_\lambda\}_{\lambda \in \Lambda}$ of A -modules, we define addition and an A -operation as follows: $\{x_\lambda\} + \{y_\lambda\} = \{x_\lambda + y_\lambda\}$, $a\{x_\lambda\} = \{ax_\lambda\}$. Then P forms

an A -module. We call $\prod_{\lambda \in \Lambda} M_\lambda$ the **direct product** of modules $\{M_\lambda\}_{\lambda \in \Lambda}$. The **canonical projection** assigning x_λ to $\{x_\lambda\}$ is denoted by $p_\lambda: P \rightarrow M_\lambda$. Suppose that an A -module M and A -homomorphisms $f_\lambda: M \rightarrow M_\lambda (\lambda \in \Lambda)$ are given. Then there exists a unique A -homomorphism $f: M \rightarrow P$ such that $p_\lambda \circ f = f_\lambda (\lambda \in \Lambda)$; f is given by $f(x) = \{f_\lambda(x)\}$.

In the direct product $\prod_{\lambda \in \Lambda} M_\lambda$, the set S of all elements whose components x_λ are equal to 0 except for a finite number of λ is called the **direct sum** of modules $\{M_\lambda\}_{\lambda \in \Lambda}$ and is denoted by $\sum_{\lambda \in \Lambda} M_\lambda$ (or $\coprod_{\lambda \in \Lambda} M_\lambda$ or $\bigoplus_{\lambda \in \Lambda} M_\lambda$). The **canonical injection** assigning $\{\dots, 0, x_\lambda, 0, \dots\} \in S$ to $x_\lambda \in M_\lambda$ is denoted by $j_\lambda: M_\lambda \rightarrow S$. If an A -module M and A -homomorphisms $f_\lambda: M_\lambda \rightarrow M$ are given, then there exists a unique A -homomorphism $f: S \rightarrow M$ such that $f \circ j_\lambda = f_\lambda (\lambda \in \Lambda)$, defined by $f(\{x_\lambda\}) = \sum_{\lambda \in \Lambda} f_\lambda(x_\lambda)$. When M is an A -module and $\{N_\lambda\}_{\lambda \in \Lambda}$ is a family of A -submodules of M , the A -homomorphism $f: \sum_{\lambda \in \Lambda} N_\lambda \rightarrow M$ defined by $f(\{x_\lambda\}) = \sum_{\lambda \in \Lambda} x_\lambda$ is an A -isomorphism if and only if M is the direct sum of $\{N_\lambda\}$.

If $M_\lambda = M$ for all $\lambda \in \Lambda$, $\prod_{\lambda \in \Lambda} M_\lambda$ and $\sum_{\lambda \in \Lambda} M_\lambda$ are denoted by M^Λ and $M^{(\Lambda)}$, respectively. M^Λ can be regarded as the set of all mappings of Λ to M . The direct product $M_1 \times \dots \times M_n$ and direct sum $M_1 \oplus \dots \oplus M_n$ of a finite number of A -modules M_1, \dots, M_n , can be identified with each other and, if $M_i = M (1 \leq i \leq n)$, we simply denote it by M^n .

G. Free Modules

Let A be a ring. A family $\{x_\lambda\}_{\lambda \in \Lambda}$ of elements in an A -module M is called **linearly independent** if $\sum_{\lambda \in \Lambda} a_\lambda x_\lambda = 0 (a_\lambda \in A)$ implies $a_\lambda = 0$ for all $\lambda \in \Lambda$. This is equivalent to saying that the mapping $A^{(\Lambda)} \rightarrow M$ that assigns $\sum_{\lambda \in \Lambda} a_\lambda x_\lambda \in M$ to $\{a_\lambda\}$ is injective. A linearly independent family $\{x_\lambda\}_{\lambda \in \Lambda}$ generating M is called a **basis** of M . A family $\{x_\lambda\}_{\lambda \in \Lambda}$ is a basis if and only if every element of M can be written uniquely in the form $\sum_{\lambda \in \Lambda} a_\lambda x_\lambda (a_\lambda \in A)$.

An A -module that has a basis is called a **free module** over A . If A is a field (which may be noncommutative), every A -module is a free module (\rightarrow 256 Linear Spaces). The \dagger cardinality of a basis of a free module M over A depends only on M if A is a field (which may be noncommutative) or a commutative ring; this number is called the **rank** (or **dimension**) of M . Any submodule of a free module over a \dagger principal ideal domain is a free module.

H. Simple Modules and Semisimple Modules

An A -module M is called **simple** if $M \neq 0$ and M has no A -submodules except M and 0.

If M and N are simple A -modules, any A -homomorphism of M to N is an isomorphism or the **zero homomorphism** (i.e., one which sends every element of M to 0) (**Schur's lemma**). If an A -module M is the sum of a family $\{M_\lambda\}_{\lambda \in \Lambda}$ of simple submodules, M is the direct sum of a suitable subfamily $\{M_{\lambda'}\}_{\lambda' \in \Lambda'}$ ($\Lambda' \subset \Lambda$). In this case, M is called **semisimple** (or **completely reducible**).

If an A -module M can be decomposed into the direct sum of A -submodules N and N' , then N' is called a **complementary submodule** of N . An A -module M is semisimple if and only if every A -submodule of M has a complementary submodule. Let A be a ring. Then the A -module A is semisimple if and only if every A -module is semisimple. In this case A is called a \dagger semisimple ring (\rightarrow 368 Rings G). Every simple module over a semisimple ring A is A -isomorphic to a \dagger minimal left ideal of A .

I. Chain Conditions

The set of all A -submodules of an A -module M forms an \dagger ordered set under the inclusion relation. An A -module is called a **Noetherian module** if the ordered set satisfies the \dagger maximal condition and an **Artinian module** if it satisfies the \dagger minimal condition (\rightarrow 311 Ordering C).

Let N be an A -submodule of an A -module M . Then M is Noetherian (Artinian) if and only if N and M/N are both Noetherian (Artinian). A ring A is called a \dagger left Noetherian ring (\dagger left Artinian ring) if A is Noetherian (Artinian) as a left A -module, and similarly for right Noetherian and Artinian rings. Every finitely generated module over a Noetherian (Artinian) ring is Noetherian (Artinian). Over an arbitrary ring A , a module M is Noetherian if and only if every A -submodule of M is finitely generated.

A finite sequence $\{M_i\}_{0 \leq i \leq r}$ of A -submodules of an A -module M is called a \dagger Jordan-Hölder sequence if $M = M_0, M_i \supset M_{i+1}, M_r = \{0\}$, and the $M_i/M_{i+1} (0 \leq i < r)$ are simple. If such a sequence exists, M is said to be of **finite length**. The number r , called the **length** of M , depends only on M . The quotient modules $M_i/M_{i+1} (0 \leq i < r)$ are uniquely determined by M up to A -isomorphism and permutation of the indices (C. Jordan and O. Hölder). An A -module M is of finite length if and only if M is Noetherian and Artinian. A semisimple A -module is of finite length if and only if it is finitely generated.

An A -module M is called **indecomposable** if M cannot be decomposed into the direct sum of two A -submodules different from M and $\{0\}$. Any A -module of finite length can be decomposed into the direct sum of a finite

sequence N_1, \dots, N_n of indecomposable A -submodules different from $\{0\}$. The direct summands N_i ($1 \leq i \leq n$) are unique up to A -isomorphism and permutation of the indices (W. Krull, R. Remak, and O. Schmidt).

J. Tensor Products

Let A be a ring. Given a right A -module M and a left A -module N , we construct a module $M \otimes_A N$ (called the **tensor product** of M and N) and a canonical mapping $\varphi: M \times N \rightarrow M \otimes_A N$ as follows. Let F be a free \mathbb{Z} -module (free Abelian additive group) generated by $M \times N$, and R be the subgroup generated by the elements of the forms $(x + x', y) - (x, y) - (x', y)$, $(x, y + y') - (x, y) - (x, y')$, $(xa, y) - (x, ay)$ ($x, x' \in M, y, y' \in N, a \in A$). We define $M \otimes_A N = F/R$, and call the natural projection φ . If we denote $\varphi(x, y)$ by $x \otimes y$, then we have $(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y$, $x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2$, and $(xa) \otimes y = x \otimes (ay)$. Any element of $M \otimes_A N$ is written in the form $\sum x_i \otimes y_i$ ($x_i \in M, y_i \in N$).

The tensor product $M \otimes_A N$ of M and N and the canonical mapping $\varphi: M \times N \rightarrow M \otimes_A N$ can be characterized as follows: For a module L , a mapping $f: M \times N \rightarrow L$ is called **biadditive** if the conditions $f(x + x', y) = f(x, y) + f(x', y)$, $f(x, y + y') = f(x, y) + f(x, y')$ hold. A biadditive mapping f satisfying the condition $f(xa, y) = f(x, ay)$ is called an **A -balanced mapping**. Then we have (i) the canonical mapping $\varphi: M \times N \rightarrow M \otimes_A N$ is A -balanced; and (ii) for any module L and any A -balanced mapping $f: M \times N \rightarrow L$, there exists a unique homomorphism $f_*: M \otimes_A N \rightarrow L$ such that $f(x, y) = f_*(x \otimes y)$ ($x \in M, y \in N$).

A right (left) A -module can be regarded as a left (right) A° -module, where A° is the ring anti-isomorphic to A . In this sense, we have $M \otimes_A N \cong N \otimes_{A^\circ} M$.

Let A be a commutative ring. For A -modules M, N , and L , a mapping $f: M \times N \rightarrow L$ is called a **bilinear mapping** if f is biadditive and satisfies $f(ax, y) = f(x, ay) = af(x, y)$ ($a \in A, x \in M, y \in N$). The set $\mathfrak{B}(M, N; L)$ of all bilinear mappings $M \times N \rightarrow L$ forms A -submodule of the A -module $L^{M \times N}$. A bilinear mapping $M \times N \rightarrow A$ is called a **bilinear form** on $M \times N$. The tensor product $M \otimes_A N$ becomes an A -module if we define $a(x \otimes y) = (ax) \otimes y (= x \otimes (ay))$, and the canonical mapping $M \times N \rightarrow M \otimes_A N$ is bilinear. For any A -module L and bilinear mapping $f: M \times N \rightarrow L$, there exists a unique A -linear mapping $f_*: M \otimes_A N \rightarrow L$ satisfying $f(x, y) = f_*(x \otimes y)$. By this correspondence $f \leftrightarrow f_*$, we get an A -isomorphism $\mathfrak{B}(M, N; L) \cong \text{Hom}_A(M \otimes_A N, L)$. If A is a field, $M \otimes_A N$ coincides with the ten-

sor product $M \otimes N$ as a linear space (\rightarrow 256 Linear Spaces H, I).

In general, let M be a B - A -bimodule and N be a left A -module. Then $M \otimes_A N$ becomes a left B -module if we define $b(x \otimes y) = (bx) \otimes y$. Let N be an A - B -bimodule and M be a right A -module. Then $M \otimes_A N$ becomes a right B -module if we define $(x \otimes y)b = x \otimes (yb)$. In particular, we have $A \otimes_A N \cong N$, $M \otimes_A A \cong M$.

Let M, M' be right A -modules and N, N' be left A -modules. For A -homomorphisms $f: M \rightarrow M'$ and $g: N \rightarrow N'$, there exists a unique homomorphism $h: M \otimes_A N \rightarrow M' \otimes_A N'$ satisfying $h(x \otimes y) = f(x) \otimes g(y)$; h is called the **tensor product** of f, g and is denoted by $f \otimes g$. We give here some simple examples (also \rightarrow Section L).

Examples. (1) Let M, N be free modules (linear spaces, for example) over a commutative ring A . If $\{x_i\}_{i \in I}$ and $\{y_j\}_{j \in J}$ are bases of M and N , respectively, $M \otimes_A N$ is also a free module with a basis $\{x_i \otimes y_j\}_{i \in I, j \in J}$. If the dimensions $\dim M, \dim N$ are finite, $\dim M \otimes_A N = \dim M \dim N$.

(2) For an \mathfrak{a} -ideal of a commutative ring A , the \mathfrak{a} -factor ring $M = A/\mathfrak{a}$ can be regarded as an A -module, and we have $M \otimes_A N \cong N/\mathfrak{a}N$. For instance, $(\mathbb{Z}/m\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/n\mathbb{Z}) \cong \mathbb{Z}/(m, n)\mathbb{Z}$, where (m, n) denotes the greatest common divisor of m and n .

K. Hom and \otimes

We continue to consider modules over a ring A . Concerning the direct sum and product, we have

$$\text{Hom}_A\left(\sum_{\lambda} M_{\lambda}, \prod_{\mu} N_{\mu}\right) \cong \prod_{\lambda, \mu} \text{Hom}_A(M_{\lambda}, N_{\mu})$$

and

$$\left(\sum_{\lambda} M_{\lambda}\right) \otimes_A \left(\sum_{\mu} N_{\mu}\right) \cong \sum_{\lambda, \mu} (M_{\lambda} \otimes_A N_{\mu}).$$

Concerning projective and inductive limits we have

$$\text{Hom}_A\left(\lim_{\rightarrow} M_{\lambda}, \lim_{\leftarrow} N_{\mu}\right) \cong \lim_{\leftarrow} \text{Hom}_A(M_{\lambda}, N_{\mu})$$

and

$$\left(\lim_{\rightarrow} M_{\lambda}\right) \otimes_A \left(\lim_{\rightarrow} N_{\mu}\right) = \lim_{\rightarrow} (M_{\lambda} \otimes_A N_{\mu}).$$

An A -homomorphism $f: M \rightarrow M'$ induces a homomorphism $\text{Hom}_A(M', N) \rightarrow \text{Hom}_A(M, N)$ by the assignment $g \rightarrow g \circ f$. An exact sequence $M' \rightarrow M \rightarrow M'' \rightarrow 0$ gives rise to the exact sequence

$$0 \rightarrow \text{Hom}_A(M'', N) \rightarrow \text{Hom}_A(M, N) \rightarrow \text{Hom}_A(M', N). \quad (1)$$

An A -homomorphism $f: N \rightarrow N'$ induces a homomorphism $\text{Hom}_A(M, N) \rightarrow \text{Hom}_A(M, N')$ by the assignments $g \rightarrow f \circ g$, and an exact sequence $0 \rightarrow N' \rightarrow N \rightarrow N''$ gives rise to the exact sequence

$$0 \rightarrow \text{Hom}_A(M, N') \rightarrow \text{Hom}_A(M, N) \rightarrow \text{Hom}_A(M, N''). \quad (2)$$

Let M be a right A -module and N', N, N'' be left A -modules. An A -homomorphism $f: N \rightarrow N'$ induces the homomorphism $1_M \otimes f: M \otimes_A N \rightarrow M \otimes_A N'$, and an exact sequence $N' \rightarrow N \rightarrow N'' \rightarrow 0$ gives rise to the exact sequence

$$M \otimes_A N' \rightarrow M \otimes_A N \rightarrow M \otimes_A N'' \rightarrow 0. \quad (3)$$

Exchanging left and right, we obtain similar results (\rightarrow 52 Categories and Functors B; 200 Homological Algebra).

Let Q be an A -module. If for any exact sequence of A -modules

$$0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0, \quad (4)$$

the induced sequence

$$0 \rightarrow \text{Hom}_A(M'', Q) \rightarrow \text{Hom}_A(M, Q) \rightarrow \text{Hom}_A(M', Q) \rightarrow 0 \quad (5)$$

is exact, then Q is called an **injective A -module**. This is equivalent to the condition that if M' is an A -submodule of an A -module M , then any A -homomorphism $M' \rightarrow Q$ can be extended to an A -homomorphism $M \rightarrow Q$. Any A -module is an A -submodule of some injective A -module, and any injective A -module is a divisible A -module. If A is a \dagger Dedekind domain, any divisible A -module is an injective A -module.

Let P be an A -module. If for any exact sequence (4), the induced sequence

$$0 \rightarrow \text{Hom}_A(P, M') \rightarrow \text{Hom}_A(P, M) \rightarrow \text{Hom}_A(P, M'') \rightarrow 0 \quad (6)$$

is exact, then P is called a **projective A -module**. This is equivalent to the condition that for any surjective A -homomorphism $g: M \rightarrow M''$ and any A -homomorphism $f: P \rightarrow M''$, there exists an A -homomorphism $h: P \rightarrow M$ satisfying $g \circ h = f$. Any A -module is a factor A -module of some projective A -module. A projective A -module has no torsion element. A free A -module is a projective A -module. In general, an A -module is a projective A -module if and only if it is a direct summand of a free A -module.

Let R be a right A -module. If for any exact sequence (4), the induced sequence

$$0 \rightarrow R \otimes_A M' \rightarrow R \otimes_A M \rightarrow R \otimes_A M'' \rightarrow 0 \quad (7)$$

is exact, then R is called a **flat A -module**. Any projective A -module is a flat A -module. A flat A -module R is called **faithfully flat** if $R \otimes_A M$

$= \{0\}$ implies $M = \{0\}$. A flat right A -module R is faithfully flat if and only if $R \neq R\mathfrak{A}$ for any left ideal \mathfrak{A} ($\neq A$) of A . Let A be a \dagger principal ideal domain. Then an A -module R is flat if and only if R has no torsion element, and R is faithfully flat if and only if R has no torsion element and $R \neq Rp$ for any \dagger prime element p of A . We have the following important examples:

(1) For a commutative ring A and its multiplicatively closed subset S , the \dagger ring of quotients A_S is flat as an A -module. However, A_S is not faithfully flat. For instance, the field of rational numbers \mathbb{Q} is not faithfully flat as a \mathbb{Z} -module.

(2) Let A be a \dagger semilocal ring and \bar{A} be its completion. Then \bar{A} is faithfully flat as an A -module (\rightarrow 284 Noetherian Rings; also [1, 7]).

In the exact sequence (4), if $\text{Im } \varphi = \text{Ker } \psi$ is a direct summand of the A -module M , we say that (4) **splits**. Then (5), (6), and (7) are exact for any A -modules Q, P, R . The exact sequence (4) splits if M' is injective or M'' is projective.

By ${}_A M, M_A$, and ${}_A M_B$, we mean that M is a left A -module, a right A -module, and an A - B -bimodule, respectively. As already stated, ${}_A M_B$ and ${}_A N$ imply ${}_B(\text{Hom}_A(M, N))$, and ${}_A M$ and ${}_A N_B$ imply $(\text{Hom}_A(M, N))_B$. Similarly ${}_B M_A$ and N_A imply $(\text{Hom}_A(M, N))_B$, and M_A and ${}_B N_A$ imply ${}_B(\text{Hom}_A(M, N))$. Furthermore, ${}_B M_A$ and ${}_A N$ imply ${}_B(M \otimes_A N)$, and M_A and ${}_A N_B$ imply $(M \otimes_A N)_B$.

With ${}_B L_A, {}_A M$, and ${}_B N$, we have

$$\text{Hom}_A(M, \text{Hom}_B(L, N)) \cong \text{Hom}_B(L \otimes_A M, N). \quad (8)$$

Similarly, for ${}_A M_B, L_A$, and N_B , we have

$$\text{Hom}_A(L, \text{Hom}_B(M, N)) \cong \text{Hom}_B(L \otimes_A M, N). \quad (8')$$

If B is a commutative ring, (8) and (8') are B -isomorphisms. Furthermore, with $L_A, {}_A M_B$, and ${}_B N$, we have

$$(L \otimes_A M) \otimes_B N \cong L \otimes_A (M \otimes_B N). \quad (9)$$

We denote by M^* the set $\text{Hom}_A(M, A)$ of all linear forms on an A -module M . Then ${}_A M$ implies M_A^* , and M_A implies ${}_A M^*$; the A -module M^* is called the **dual module** of M . A as a left A -module is dual to A as a right A -module, and vice versa. For a family of A -modules $\{M_\lambda\}_{\lambda \in \Lambda}$, we have a canonical correspondence $(\sum_{\lambda \in \Lambda} M_\lambda)^* \cong \prod_{\lambda \in \Lambda} M_\lambda^*$. From this, we have a canonical isomorphism $(M^*)^* \cong M$ for any finitely generated projective A -module M . Many facts concerning this \dagger duality are similar to those valid for linear spaces (\rightarrow 256 Linear Spaces G).

Let A be a commutative ring. Letting $A = B = N$ in (8) and (8'), we have the canonical A -

isomorphisms

$$\text{Hom}_A(M, L^*) \cong \text{Hom}_A(L, M^*)$$

$$\cong (L \otimes_A M)^* = \mathfrak{L}(L, M; A),$$

namely, any bilinear form on $L \times M$ is represented by a linear mapping $M \rightarrow L^*$ or $L \rightarrow M^*$.

L. Extension and Restriction of a Basic Ring

Fix a ring homomorphism $\rho: A \rightarrow B$. We regard B as a B - A -bimodule by defining a right operation of A on B by $b \cdot a = b\rho(a)$ ($a \in A, b \in B$). This bimodule is denoted by B_ρ .

For every left A -module M , we construct the left B -module $\rho^*(M) = B_\rho \otimes_A M$, which is called the **scalar extension** of M by ρ . Every A -homomorphism of A -modules $f: M \rightarrow M'$ induces the B -homomorphism $\rho^*(f) = 1_B \otimes f: \rho^*(M) \rightarrow \rho^*(M')$.

For every left B -module N , we construct the left A -module $\rho_*(N) = \text{Hom}_B(B_\rho, N)$, which is called the **scalar restriction** (or **scalar change**) of N by ρ . By the assignment $h \rightarrow h(1)$, we have a module isomorphism $\text{Hom}_B(B_\rho, N) \cong N$. We identify $\rho_*(N)$ with N under this isomorphism. The operation of A on N is then given as $a \cdot y = \rho(a)y$ ($a \in A, y \in N$). If A is a subring of B and ρ is the canonical injection, then the operation of A on $\rho_*(N)$ is the restriction of the operation of B on N . Every B -homomorphism of B -modules $f: N \rightarrow N'$ induces the A -homomorphism $\rho_*(f): \rho_*(N) \rightarrow \rho_*(N')$. For any left A -module M and left B -module N , an A -linear mapping $f: M \rightarrow \rho_*(N) = N$ is called a **semilinear mapping** with respect to ρ . This means that f is an additive homomorphism satisfying $f(ax) = \rho(a)f(x)$ ($a \in A, x \in M$).

The extension and the restriction of a basic ring are related by the canonical isomorphism $\text{Hom}_A(M, \rho_*(N)) \cong \text{Hom}_B(\rho^*(M), N)$ for an A -module M and a B -module N (\rightarrow equation (8')). An element α of the left-hand side and an element β of the right-hand side are associated by the relation $\alpha(x) = \beta(1 \otimes x)$ ($x \in M$) (\rightarrow 52 Categories and Functors).

Let A and B be commutative rings. Then for A -modules M and M' , we have the canonical B -linear mapping $\rho^*: B \otimes_A \text{Hom}_A(M, M') \rightarrow \text{Hom}_B(B \otimes_A M, B \otimes_A M')$, which is a B -isomorphism if M is a finitely generated free (or more generally, projective) A -module. Using the notation ρ^* , we have $\rho^*(\text{Hom}_A(M, M')) \cong \text{Hom}_B(\rho^*(M), \rho^*(M'))$.

We now give some examples where the basic rings are noncommutative. Let G be a group and H its subgroup. Let ρ denote the homomorphism of group rings $K[H] \rightarrow K[G]$ induced by the canonical injection $H \rightarrow G$, where K is a commutative ring. For any $K[H]$ -module M , $\rho^*(M)$ is called the **induced module**

of M . The representation of G associated with $\rho^*(M)$ is the \dagger induced representation of the representation of H associated with M . Next, we fix a group G and consider a homomorphism $\rho: K[G] \rightarrow \bar{K}[G]$ induced by a homomorphism of commutative rings $\sigma: K \rightarrow \bar{K}$. If $K = \bar{K}$ and σ is an automorphism, then the representation associated with the “scalar extension” $\rho^*(M)$ of a $K[G]$ -module M is the \dagger conjugate representation to the representation associated with M . If $\bar{K} = K/\mathfrak{A}$ (\mathfrak{A} is an ideal of K) and σ is the canonical projection, then the representation over \bar{K} associated with the scalar extension $\rho^*(M)$ of a $K[G]$ -module M is the **reduction modulo \mathfrak{A}** of the representation over K associated with M , and $\rho^*(M)$ is canonically isomorphic to $M/\mathfrak{A}M$. Furthermore, the \dagger localization and the \dagger completion can also be treated under the formulation of scalar extension (\rightarrow 67 Commutative Rings G, 284 Noetherian Rings B).

References

- [1] M. Nagata, Local rings, Interscience, 1962.
- [2] H. P. Cartan and S. Eilenberg, Homological algebra, Princeton Univ. Press, 1956.
- [3] D. G. Northcott, An introduction to homological algebra, Cambridge Univ. Press, 1960.
- [4] S. MacLane, Homology, Springer, 1963.
- [5] C. Chevalley, Fundamental concepts of algebra, Academic Press, 1956.
- [6] N. Bourbaki, Eléments de mathématique, Algèbre, ch. 2, Actualités Sci. Ind., 1236b, Hermann, third edition, 1962.
- [7] N. Bourbaki, Eléments de mathématique, Algèbre commutative, ch. 1, 2, Actualités Sci. Ind., 1290a, Hermann, 1961.
- [8] C. W. Curtis and I. Reiner, Representation theory of finite groups and associative algebras, Interscience, 1962.
- [9] R. Godement, Cours d'algèbre, Hermann, 1963.
- [10] S. Lang, Algebra, Addison-Wesley, 1965.
- [11] S. T. Hu, Elements of modern algebra, Holden-Day, 1965.

278 (XIII.23) Monge-Ampère Equations

A. Monge-Ampère Equations

A **Monge-Ampère differential equation** is a second-order partial differential equation of the form

$$Hr + 2Ks + Lt + M + N(rt - s^2) = 0, \quad (1)$$

where H, K, L, M , and N are functions of x, y, z, p , and q , and r, s, t, p , and q represent the partial derivatives

$$r = \frac{\partial^2 z}{\partial x^2}, \quad s = \frac{\partial^2 z}{\partial x \partial y}, \quad t = \frac{\partial^2 z}{\partial y^2},$$

$$p = \frac{\partial z}{\partial x}, \quad q = \frac{\partial z}{\partial y}.$$

The characteristic manifolds are integrals of a system of differential equations defined as follows:

Case (i) $N \neq 0$.

$$\begin{aligned} N dp + L dx + \lambda_1 dy &= 0, \\ N dq + \lambda_2 dx + H dy &= 0, \\ dz - p dx - q dy &= 0, \end{aligned} \quad (2)$$

$$\begin{aligned} N dp + L dx + \lambda_2 dy &= 0, \\ N dq + \lambda_1 dx + H dy &= 0, \\ dz - p dx - q dy &= 0, \end{aligned} \quad (3)$$

where λ_1 and λ_2 are the two roots of the equation $\lambda^2 + 2K\lambda + HL - MN = 0$.

Case (ii) $N = 0, H \neq 0$.

$$\begin{aligned} dy &= \lambda_1 dx, \quad H dp + H \lambda_2 dq + M dx = 0, \\ dz - p dx - q dy &= 0, \end{aligned} \quad (4)$$

$$\begin{aligned} dy &= \lambda_2 dx, \quad H dp + H \lambda_1 dq + M dx = 0, \\ dz - p dx - q dy &= 0, \end{aligned} \quad (5)$$

where λ_1 and λ_2 are the two roots of the equation $H\lambda^2 - 2K\lambda + L = 0$.

Case (iii) $N = 0, H = 0, L \neq 0$.

$$\begin{aligned} dx &= 0, \quad M dy + 2K dp + L dq = 0, \\ dz - p dx - q dy &= 0, \end{aligned} \quad (6)$$

$$\begin{aligned} 2K dy - L dx &= 0, \quad M dy + L dq = 0, \\ dz - p dx - q dy &= 0. \end{aligned} \quad (7)$$

Case (iv) $N = 0, H = 0, L = 0$.

$$\begin{aligned} dx &= 0, \quad 2K dp + M dy = 0, \\ dz - p dx - q dy &= 0, \end{aligned} \quad (8)$$

$$\begin{aligned} dy &= 0, \quad 2K dq + M dx = 0, \\ dz - p dx - q dy &= 0. \end{aligned} \quad (9)$$

A manifold $x(\lambda), y(\lambda), z(\lambda), p(\lambda), q(\lambda)$ that satisfies the system (2), (3) of differential equations for case (i), (4), (5) for case (ii), (6), (7) for case (iii), or (8), (9) for case (iv) is a characteristic manifold of equation (1).

The following result is known concerning Monge-Ampère equations: The union of surface elements of an integral surface of (1) is generated in two ways by characteristic manifolds depending on one parameter. Conversely,

if a manifold, each element of which is composed of a point of a surface S and the tangent plane at that point, is generated by a family of characteristic manifolds depending on one parameter, then the surface S is an integral surface of (1).

B. Intermediate Integrals

If a relation $dV(x, y, z, p, q) = 0$ is a consequence of a system of differential equations of characteristic manifolds, for example, in case (i) when V satisfies

$$\begin{aligned} N(\partial V / \partial x + p \partial V / \partial z) - L \partial V / \partial p - \lambda_2 \partial V / \partial q &= 0, \\ N(\partial V / \partial y + q \partial V / \partial z) - \lambda_1 \partial V / \partial p - H \partial V / \partial q &= 0, \end{aligned}$$

$V(x, y, z, p, q) = c$ (c an arbitrary constant) is called an **integral** of the system of differential equations. (i) If $V = c$ is an integral of a system of differential equations of characteristic manifolds, the solution $z(x, y)$ of $V = c$ considered anew as a partial differential equation of the first order is a solution of (1). Conversely, if every solution of $V = c$ (excepting *singular ones) satisfies equation (1), $V = c$ is an integral of a system of differential equations of characteristic manifolds. (ii) If $u(x, y, z, p, q)$ and $v(x, y, z, p, q)$ are two integrals of a system of differential equations of characteristic manifolds, then for any arbitrary function φ , $\varphi(u, v)$ is also an integral. Thus, as a consequence of (i), every solution $z(x, y)$ of $\varphi(u, v) = 0$ is also a solution of (1). The converse is also true, namely, for every solution z of (1) we can find a function φ such that $z(x, y)$ is also a solution of $\varphi(u, v) = 0$. The relation $\varphi(u, v) = 0$ is called an **intermediate integral** of (1). Sometimes an integral of a system of differential equations of characteristic manifolds is also called an intermediate integral. If each of the two systems of differential equations defining the characteristic manifolds has an intermediate integral, then the two intermediate integrals $\varphi(u, v) = 0$ and $\psi(u, v) = 0$ form a *complete system of partial differential equations of the first order. Integrating this complete system, the *general solution of equation (1) is obtained.

References

- [1] E. Goursat, Cours d'analyse mathématique III, Gauthier-Villars, fourth edition, 1927.
- [2] E. Goursat, Leçons sur l'intégration des équations aux dérivées partielles du second ordre à deux variables indépendantes I, Hermann, 1896.

279 (VII.16) Morse Theory

A. General Remarks

We are interested in smooth functions on a smooth manifold of dimension m that have only the simplest critical points. Here, "smooth" means differentiable of class C^∞ .

Such a function $f: M \rightarrow \mathbf{R}$ enables us to investigate the topology of M . The decomposition of M into level sets of f contains a lot of topological information on M . For instance, it is shown that M is homotopy equivalent to a CW complex that is determined by critical points of f ; furthermore the Euler characteristic of M can be computed by means of f , because the critical points are related to the homology groups of M . These types of investigations, called **Morse theory**, were originated by H. Poincaré [1] and G. D. Birkhoff [2] and then developed into the form we see today by M. Morse [3]. An excellent exposition of this theory has been given by J. Milnor [4]. R. S. Palais [5] has extended the theory to Hilbert manifolds. Morse theory has been fruitfully applied to differential topology and differential geometry.

B. Critical Points of Functions on Manifolds

For a point $p \in M$, M_p is by definition the tangent space to M at p . Let f be a smooth real-valued function on M . A point $p \in M$ is called a **critical point** of f if the induced mapping $f_*: M_p \rightarrow \mathbf{R}_{f(p)}$ is zero at p . For every local chart (x_1, \dots, x_m) around a critical point p of f with $p = (0, \dots, 0)$,

$$\frac{\partial f}{\partial x_1}(0) = \dots = \frac{\partial f}{\partial x_m}(0) = 0.$$

The value $f(p)$ is then called a **critical value** of f . If the matrix $(\partial^2 f / \partial x_i \partial x_j(0))$ is invertible, then the point is called a **nondegenerate critical point**, and if the matrix is not invertible, then the point is said to be **degenerate**. These notions are independent of the choice of local charts around p . The above matrix is called the **Hessian** of f at p . The nullity and index of the Hessian of f at p are called the **nullity** and **index** of f at p , respectively. The function f has a local minimum at a nondegenerate critical point of index 0, and a local maximum at a nondegenerate critical point of index n .

A smooth function f on M is called a **Morse function** if it satisfies: (A1) Every critical point of f is nondegenerate. The **Morse lemma** states that if p is a nondegenerate critical point of f

and if the index of f at p is λ , then there exists a local chart (y_1, \dots, y_m) around p with $p = (0, \dots, 0)$ such that f is expressed as

$$f(y) = f(p) - (y_1)^2 - \dots - (y_\lambda)^2 + (y_{\lambda+1})^2 + \dots + (y_m)^2.$$

C. The Existence of Morse Functions

Let M be a compact smooth m -dimensional manifold without boundary. The existence of Morse functions on M is shown by embedding M into \mathbf{R}^n for a sufficiently large n . For convenience, M will be here identified with its embedded image. At each point $p \in M$ the set v_p of all unit normals forms an $(n-m-1)$ -dimensional unit sphere, and $v(M) = \bigcup_{p \in M} v_p$ is a smooth $(n-1)$ -dimensional compact manifold. Let S be the unit hypersphere in \mathbf{R}^n around the origin. The Gauss normal mapping $\varphi: v(M) \rightarrow S$ sends each point $(p, v(p)) \in v(M)$ to $v \in S$ via a canonical parallel translation of \mathbf{R}^n . Since φ is smooth, the Sard theorem implies that the set E of all critical values of φ has measure zero on S . Thus φ_* at every $(p, v(p)) \in \varphi^{-1}(v)$ has maximal rank if v is taken in $S - E$. For every fixed $v \in S - E$ let $h_v: M \rightarrow \mathbf{R}$ be defined by $h_v(x) = \langle v, x \rangle$, $x \in M$, where $\langle \cdot, \cdot \rangle$ denotes the canonical inner product of \mathbf{R}^n . Then h_v is a Morse function. To see this let dS , dv , and dM be volume elements of S , v , and M , respectively. Then $dv \wedge dM$ is a volume element of $v(M)$, and the function $G: v(M) \rightarrow \mathbf{R}$ is obtained by the relation $\varphi^* dS = G(x, v(x)) dv \wedge dM$. $G(x, v(x))$ is called the **Lipschitz-Killing curvature** at $v(x) \in v_x$. If $A(v(x))$ denotes the second fundamental form of M with respect to $v(x)$, then $G(x, v(x)) = (-1)^m \det(A(v(x)))$ (Chern and Lashof [6]). Every critical point $p \in M$ of h_v for $v \in S - E$ has a normal $v(p)$ to M at p , and the Hessian of h_v is given by $A(v(p))$. Thus $G(p, v(p)) \neq 0$ ensures the nondegeneracy at p [6].

Another approach to the construction of Morse functions is to find for a given embedding of M into \mathbf{R}^n an open dense set $U \subset \mathbf{R}^n$ such that the Euclidean distance d_x from any fixed point x on U to points on M has no degenerate critical points on M . In this case $d_x^{-1}((-\infty, a])$ is compact for all $a \in \mathbf{R}$. It is also possible to construct Morse functions on compact manifolds with boundary as well as on noncompact manifolds. But a striking result of A. Phillips states that there exists a Morse function on every noncompact manifold which has no critical points [7].

For a pair of smooth manifolds M, N let $C^\infty(M, N)$ be the set of all smooth mappings from M to N . The **weak** (or **compact-open**) C^∞

topology on $C^\infty(M, N)$ is generated by the sets defined as follows. Let $f \in C^\infty(M, N)$, and let (U, φ) and (V, ψ) be charts of M and N , respectively. Let $K \subset U$ be a compact set such that $f(K) \subset V$. For $\varepsilon \in (0, \infty)$, a weak \dagger subbasic neighborhood $N^\infty(f; (U, \varphi), (V, \psi), K, \varepsilon)$ of f is defined to be the set of all smooth mappings $g \in C^\infty(M, N)$ such that (1) $g(K) \subset V$, (2) $\|D^k(\varphi \circ f \circ \psi^{-1})(x) - D^k(\varphi \circ g \circ \psi^{-1})(x)\| < \varepsilon$, $x \in \varphi(K)$, $k = 1, 2, \dots$, where D^k denotes the k th derivative with respect to the coordinates. The weak C^∞ topology is thus defined on $C^\infty(M, N)$, and the space with this topology is denoted by $C_w^\infty(M, N)$. It follows from the construction of the embedding in the Whitney embedding theorem that the set of all smooth embeddings of compact M into \mathbf{R}^n is dense in $C_w^\infty(M, \mathbf{R}^n)$ if $n > 2 \dim M$. Moreover, the set of all Morse functions on a compact M forms an open dense set in $C_w^\infty(M, \mathbf{R})$ (Nagano [8], Hirsh [9], Auslander and MacKenzie [10]).

D. Decomposition of M by a Morse Function

To decompose M into levels of a Morse function we now consider the process of \dagger attaching a handle. Let M be a compact manifold with \dagger boundary ∂M . Let D^s be the s -disk, and let $g: (\partial D^s) \times D^{m-s} \rightarrow \partial M$ be an \dagger embedding. Then a \dagger manifold $X(M; g; s)$ with a handle attached by g is defined as the quotient set obtained from the disjoint union $M \cup (D^s \times D^{m-s})$ by identifying points in $\partial D^s \times D^{m-s}$ and their images under g and equipped with a natural differentiable structure. Similarly, if $g_i: (\partial D_i^s) \times D_i^{m-s_i} \rightarrow \partial M$ ($i = 1, \dots, k$) are embeddings with disjoint images, we can define the \dagger handlebody $X(M; g_1, \dots, g_k; s_1, \dots, s_k)$ [11].

For an $f \in C^\infty(M, \mathbf{R})$ and for $a, b \in f(M)$ with $a < b$ let $M^a = \{x \in M \mid f(x) \leq a\}$ and $M_a^b = \{x \in M \mid a \leq f(x) \leq b\}$. M_a^b is called a **level set** of f . For a Morse function f on a compact manifold M without boundary, the following fundamental facts are known [4]: (1) The **first fundamental theorem**. If M_a^b contains no critical points, then M_a^b is diffeomorphic to $M_a^a \times [a, b]$; (2) The **second fundamental theorem**. Let $c \in (a, b)$ be a unique critical value in $[a, b]$, and let $p_1, \dots, p_k \in M_c^c$ be critical points of f with p_i having index λ_i . Then M^b is diffeomorphic to $X(M^a; f_1, \dots, f_k, \lambda_1, \dots, \lambda_k)$ for suitable embeddings f_1, \dots, f_k . It follows from the existence of Morse functions and from (1) and (2) that every compact manifold can be obtained by successively attaching handles to a disk (\rightarrow 114 Differential Topology). Furthermore, if M admits a Morse function with only two critical points, then M is homeomorphic to a sphere (Reeb [12]).

Concerning the homotopy types of M and

M^a we have the following: (3) If $c \in f(M)$ is a critical value of f with points p_1, \dots, p_k with p_i having index λ_i and if $M_c^{c+\varepsilon}$ is compact for $\varepsilon > 0$ and contains no critical points other than p_1, \dots, p_k , then $M^{c+\varepsilon}$ has the same homotopy type as $M^{c-\varepsilon} \cup e^{\lambda_1} \cup \dots \cup e^{\lambda_k}$, where e^{λ_j} ($j = 1, \dots, k$) are λ_j -cells. (4) Assume that a Morse function f on a (not necessarily compact) M satisfies: (A2) M^a is compact for all $a \in \mathbf{R}$. Then by choosing a sequence $a_1 < a_2 < \dots$ of regular values of f and applying (3) to each $M_{a_i}^{a_{i+1}}$, we see that M has the same homotopy type as that of a CW complex ζ , where the number of λ -cells belonging to ζ is the same as that of critical points of index λ of f .

Moreover, we have the following Morse inequalities: (5) Let f be a Morse function on a compact M . If M_λ is the number of critical points of f on M of index λ , and if R_λ is a λ -dimensional \dagger Betti number of M , then

$$M_0 \geq R_0,$$

$$M_1 - M_0 \geq R_1 - R_0,$$

$$M_\lambda - M_{\lambda-1} + \dots + (-1)^\lambda M_0$$

$$\geq R_\lambda - R_{\lambda-1} + \dots + (-1)^\lambda R_0, \quad 1 \leq \lambda \leq n-1,$$

$$M_n - M_{n-1} + \dots + (-1)^n M_0$$

$$= R_n - R_{n-1} + \dots + (-1)^n R_0.$$

In particular, we have $M_k \geq R_k$ for all k . By using these facts S. Smale obtained an affirmative solution of the \dagger Poincaré conjecture in high dimensions [11, 13].

The concept of critical manifolds in the sense of Bott [14] is stated as follows: Let M be a compact manifold embedded into an open set $U \subset \mathbf{R}^k$. Let $f: U \rightarrow \mathbf{R}$ be a smooth function. M is called a **nondegenerate critical manifold** of f on U if (1) all points of M are critical points of f and (2) the nullity of all $x \in M$ is equal to $\dim M$. When M is such a nondegenerate critical manifold of f on U , f is constant on M , and the index λ of f at x is well defined and is the same at all points on M . The \dagger Poincaré polynomial is expressed as $P(M; t) = \sum t^k \dim H^k(M)$. The Morse polynomial is defined by $\mathfrak{M}(f; t) = \sum_N t^{\lambda_N} P(N; t)$, where N runs over all critical manifolds of f and λ_N is the index of N . Under certain conditions for orientability along each nondegenerate critical manifold of f , we have again the **Morse inequality**

$$\mathfrak{M}(f; t) \geq P(M; t).$$

E. Morse Theory on Hilbert Manifolds

Let M be a \dagger Hilbert manifold, and let $f: M \rightarrow \mathbf{R}$ be a smooth function. The Hessian $\partial^2 f_p$ of f at a critical point of f is a symmetric bilinear

form on M_p given by

$$\partial^2 f_p(u, v) = d^2(f \circ \varphi^{-1})(d\varphi_p(u), d\varphi_p(v)),$$
$$u, v \in M_p,$$

where φ is a coordinate mapping of a local chart around p . The **index (coindex)** of f at a critical point $p \in M$ of f is defined to be the supremum of dimensions of subspaces of M_p on which $\partial^2 f_p$ is negative (positive) definite. The self-adjoint bounded operator A which represents $\partial^2 f_p$ is given by $\partial^2 f_p(u, v) = \langle u, A(v) \rangle_p$. If A is invertible, then f is said to be **nondegenerate** at p . These notions do not depend on the choice of local charts around p .

The Morse lemma has been generalized to a Riemannian Hilbert manifold M by R. Palais and S. Smale [15] as follows. Let f be a smooth function defined on M . If $x \in M$ is a nondegenerate critical point of f then there exist a chart φ around x and a projection P such that for y near x

$$f(y) = f(x) + \|P\varphi(y)\|^2 - \|(I - P)\varphi(y)\|^2.$$

The above result has been extended to critical manifolds [16]. A connected submanifold N of a Hilbert manifold M is called a **nondegenerate critical manifold** of $f: M \rightarrow \mathbf{R}$ if (1) every point $p \in M$ is a critical point of f and (2) for each $p \in M$ there exists a closed subspace E_p of M_p such that $M_p = N_p \oplus E_p$ and the Hessian form restricted to E_p is nondegenerate. Let $v = (\pi, E, N)$ be a smooth Hilbert-space bundle over a compact connected manifold N , and let $\langle \cdot, \cdot \rangle$ be a Riemannian structure for v . Assume that a smooth function $f: E \rightarrow \mathbf{R}$ has the zero section of v as a nondegenerate critical manifold. Then there exist a tubular neighborhood $v_\varepsilon = \{v \in E \mid \|v\| < \varepsilon\}$ of the zero section in v and a fiber-preserving diffeomorphism $\psi: v_\varepsilon \rightarrow \psi(v_\varepsilon)$ and an orthogonal bundle projection P such that for $v \in v_\varepsilon$

$$f \circ \psi(v) = f(N) + \|Pv\|^2 - \|(I - P)v\|^2.$$

Let us assume that (B) M is a complete Riemannian manifold, and (C) (**Palais-Smale condition**) if f is bounded on a set $S \subset M$ and if the norm $\|\nabla f\|$ of the \dagger gradient vector of f has infimum 0 on S , then there is a critical point of f on the closure \bar{S} of S . Since M is not locally compact, condition (C) is required. In order to prove the first and second fundamental theorems of Morse theory it is necessary that the integral curves of ∇f exist, and (C) ensures their existence. Namely, under conditions (A1), (B), and (C), one of the following two facts follows for any $b \in f(M)$ and for any regular point $p \in M^b$: (1) The integral curve $c: [0, r] \rightarrow M^b$ of ∇f with $c(0) = p$ exists, and $f \circ c(r) = b$ holds for some $r \in [0, \infty)$; (2) the integral curve $c: [0, \infty) \rightarrow M^b$ of ∇f exists and

$\lim_{t \rightarrow \infty} c(t)$ exists such that the limit point $c(\infty) = \lim_{t \rightarrow \infty} c(t) \in M^b$ is a critical point of f . These facts imply that the first fundamental theorem of Morse theory holds if M_a^b does not contain any critical point of f , and also the second fundamental theorem holds if M_a^b has only nondegenerate critical points on M_c^c for some $c \in (a, b)$. Furthermore, assume that $f: M \rightarrow \mathbf{R}$ and M satisfies (A1), (B), and (C), and let $a, b \in f(M)$ be regular values with $a < b$. If for each nonnegative integer λ , R_λ is the λ th Betti number of M_a^b , and if M_λ^b is the number of critical points of index λ of f in M_a^b , then the **Morse inequality** holds:

$$M_0 \geq R_0,$$
$$M_1 - M_0 \geq R_1 - R_0,$$
$$\dots,$$
$$\sum_{k=0}^{\lambda} (-1)^{\lambda-k} M_k \geq \sum_{k=0}^{\lambda} (-1)^{\lambda-k} R_k,$$
$$\dots,$$
$$\sum_{k=0}^{\infty} (-1)^k M_k = \sum_{k=0}^{\infty} (-1)^k R_k.$$

In particular, $M_\lambda \geq R_\lambda$ holds for all λ . If f is bounded below, then the λ th Betti number R_λ^* of M^b and the number M_λ^* of all critical points of f with index λ in M^b satisfy $M_\lambda^* \geq R_\lambda^*$ for all λ .

F. Morse Theory of Path Spaces

Morse theory on Hilbert manifolds applies to the energy functions on Riemannian Hilbert manifolds which consist of all H_1 -curves on a compact Riemannian manifold M , and the theory is useful for proving the existence of closed geodesics on M .

Let M be a smooth manifold, and let $I = [0, 1]$. Let $H_1(I, M)$ be the set of all continuous curves $\sigma: I \rightarrow M$ such that for each local chart (V, φ) of M , $\varphi \circ \sigma$ is absolutely continuous and $\|(\varphi \circ \sigma)'\|$ is locally square integrable. In particular, if $M = \mathbf{R}^m$, then $H_1(I, \mathbf{R}^m)$ is a Hilbert space with the inner product

$$(\sigma, \rho)_1 := (\sigma(0), \rho(0)) + \int_I (\sigma(t), \rho(t)) dt,$$
$$\sigma, \rho \in H_1(I, \mathbf{R}^m).$$

For each $\sigma \in H_1(I, M)$, set $H_1(I, M)_\sigma = \{X \in H_1(I, TM) \mid X(t) \in M_{\sigma(t)}, t \in I\}$, where TM is by definition the \dagger tangent bundle over M . For any fixed pair of points $p, q \in M$, let $\Omega(M; p, q) = \{\sigma \in H_1(I, M) \mid \sigma(0) = p, \sigma(1) = q\}$, and for each $\sigma \in \Omega(M; p, q)$, let $\Omega(M; p, q)_\sigma = \{X \in H_1(I, M)_\sigma \mid X(0) = 0 \in M_p, X(1) = 0 \in M_q\}$. Then $H_1(I, M)_\sigma$ forms a vector space, and $\Omega(M; p, q)_\sigma$ is a subspace of $H_1(I, M)_\sigma$. M can be embedded into \mathbf{R}^n for sufficiently large n by the Whitney

embedding theorem [17]. Then we have the following facts [17]: (1) $H_1(I, M) = \{\sigma \in H_1(I, \mathbf{R}^n) \mid \sigma(I) \subset M\}$, (2) $H_1(I, M)$ is a closed submanifold of $H_1(I, \mathbf{R}^n)$, (3) $\Omega(M; p, q)$ is a closed submanifold of $H_1(I, M)$, (4) the tangent space to $H_1(I, M)$ at each point $\sigma \in H_1(I, M)$ is $H_1(I, M)_\sigma$, (5) the tangent space to $\Omega(M; p, q)$ at each point $\sigma \in \Omega(M; p, q)$ is $\Omega(M; p, q)_\sigma$. Here M is identified with the embedded image in \mathbf{R}^n . Thus $H_1(I, M)$ and $\Omega(M; p, q)$ carry the structure of a Hilbert manifold, and they are independent of the choice of embeddings [17].

The energy functions on $H_1(I, M)$ and $\Omega(M; p, q)$ play important roles in the development of Morse theory. M is now assumed to be a complete Riemannian manifold. It follows from the Nash isometric embedding theorem (\rightarrow 365 Riemannian Submanifolds B) that M is isometrically embedded into \mathbf{R}^n for a sufficiently large n . $H_1(I, \mathbf{R}^n)$ carries a complete Riemannian metric in a natural way, and $H_1(I, M)$ becomes a complete Riemannian manifold with the metric induced from $H_1(I, \mathbf{R}^n)$. Similarly $\Omega(M; p, q)$ admits a complete Riemannian metric. Then the **energy function** $E: H_1(I, M) \rightarrow \mathbf{R}$ is defined to be

$$E(\sigma) := \frac{1}{2} \int_I \langle \sigma', \sigma' \rangle dt,$$

where $\langle \cdot, \cdot \rangle$ is the inner product induced from the Riemannian structure of M . Then the Palais-Smale condition (C) is satisfied for E on $H_1(I, M)$. That is, if a sequence $\{\sigma_k\}$ on $H_1(I, M)$ satisfies: (1) $\{E(\sigma_k)\}$ is bounded above; and (2) $\nabla E(\sigma_k) \rightarrow 0$ as $k \rightarrow \infty$, then there is a subsequence $\{\sigma_{k'}\}$ of $\{\sigma_k\}$ such that $\{\sigma_{k'}\}$ converges to a critical point $\sigma \in H_1(I, M)$ of E . Also, condition (C) is fulfilled for the energy function on $\Omega(M; p, q)$.

For a fixed pair of points $p, q \in M$, let $\Omega = \Omega(M, p, q)$. Let $\sigma \in \Omega$ be a critical point of E . A tangent vector W to Ω at $\sigma: I \rightarrow M$ with $\sigma(0) = p, \sigma(1) = q$ is a \dagger piecewise differentiable vector field along σ such that $W(0) = 0 \in M_p$ and $W(1) = 0 \in M_q$. A **proper variation** $\alpha: (-\varepsilon, \varepsilon) \times I \rightarrow M$ along σ which is associated with W satisfies the following: (1) $\alpha(0, t) = \sigma(t), t \in I$; (2) there exists a finite partition $0 = t_0 < t_1 < \dots < t_k = 1$ of I such that $\alpha|(-\varepsilon, \varepsilon) \times [t_{i-1}, t_i]$ is differentiable for all $i = 1, \dots, k$; (3) $\alpha(u, 0) = p$ and $\alpha(u, 1) = q$ hold for all $u \in (-\varepsilon, \varepsilon)$; and (4) $\partial\alpha(0, t)/\partial u = W(t), t \in I$. It follows from the first variation formula (\rightarrow 178 Geodesics A) that $\sigma \in \Omega$ is a critical point of E if and only if σ is a geodesic on M . Let $W_1, W_2 \in \Omega_\sigma$, and let U be an open set in \mathbf{R}^2 around the origin. Then a **proper variation** $\alpha: U \times I \rightarrow M$ along σ that is associated with W_1 and W_2 satisfies the following: (1) $\alpha(0, t) = \sigma(t), t \in I$; (2) there exists a finite partition $0 = t_0 < t_1 < \dots < t_k = 1$ of I such

that $\alpha|U \times [t_{i-1}, t_i]$ is differentiable for all $i = 1, \dots, k$; (3) $\alpha(u_1, u_2, 0) = p, \alpha(u_1, u_2, 1) = q$ for all $(u_1, u_2) \in U$; and (4) $\partial\alpha(0, 0, t)/\partial u_1 = W_1(t), \partial\alpha(0, 0, t)/\partial u_2 = W_2(t), t \in I$. Then the **Hessian** E_{**} of E at σ is given by

$$E_{**}(W_1, W_2) = \frac{1}{2} \frac{\partial^2 E(\bar{\alpha}(u_1, u_2))}{\partial u_1 \partial u_2} \Big|_{(0,0)},$$

where $\bar{\alpha}(u_1, u_2) \in \Omega$ is by definition the curve $\bar{\alpha}(u_1, u_2)(t) = \alpha(u_1, u_2, t)$. The second variation formula (\rightarrow 178 Geodesics A) then gives

$$E_{**}(W_1, W_2) = - \sum \langle W_2(t), \Delta_t W_1'(t) \rangle - \int_I \langle W_2, W_1'' + R(W_1, \sigma')\sigma' \rangle dt,$$

where $\Delta_t W_1'(t) = W_1'(t+0) - W_1'(t-0)$. We have the **Morse index theorem**, which states: (1) The \dagger null space of E at a critical point σ is the linear space spanned by all Jacobi fields (\rightarrow 178 Geodesics A) along σ that vanish at 0 and 1; (2) if $\{\sigma(s_1), \sigma(s_2), \dots, \sigma(s_l)\}$ ($0 < s_1 < s_2 < \dots < s_l < 1$) is the set of all points conjugate to $\sigma(0)$ along σ and if λ_i is the multiplicity of the conjugate point $\sigma(s_i)$, then the index of E at σ is equal to $\lambda_1 + \dots + \lambda_l$. It follows from the Sard theorem together with the differentiability of the exponential mapping on M that except for a set of measure zero in $M \times M, p, q$ can be chosen so that p, q is not a conjugate pair along any geodesic in Ω . Then all critical points of E are nondegenerate, and for any $c > 0, \Omega^c = \{\omega \in \Omega \mid E(\omega) = c\}$ contains at most finitely many nondegenerate critical points with finite indices.

G. Existence of Closed Geodesics

Let M be a compact Riemannian manifold. By replacing I with the circle S^1 , we consider a Hilbert manifold $\Lambda(M) = H_1(S^1, M)$. $\Lambda(M)$ carries the structure of a complete Riemannian manifold. Every point $p \in M$ is naturally embedded in $\Lambda(M)$ as a point curve, and M is a totally geodesic submanifold of $\Lambda(M)$. A point $\sigma \in \Lambda(M)$ is a critical point of the energy function $E: \Lambda(M) \rightarrow \mathbf{R}$ if and only if σ is a closed geodesic of M . It is known that E satisfies the Palais-Smale condition (C). The index and nullity of E at a critical point σ is finite. Since M is compact, the fundamental length \sqrt{c} of M is positive and $\Lambda^c = \{\sigma \in \Lambda(M) \mid E(\sigma) \leq c\}$ is a \dagger deformation retract of $M \subset \Lambda(M)$ by means of the deformation along the integral curves of $-\nabla E$.

If M is not simply connected, then each nontrivial element of the \dagger fundamental group of M represents a class of homotopic curves in

which there is a closed geodesic whose length realizes the infimum of all these curves. If M is simply connected, then there is a minimum integer k for which $\pi_k(M) \neq 0$. Clearly, $\pi_{k-1}(\Lambda(M)) \approx \pi_k(M) \neq 0$. Suppose that E has no positive critical value. Then $\Lambda(M)$ is a deformation retract of M , a contradiction. Therefore there exists at least one closed geodesic on every compact Riemannian manifold (Lyusternik and Fet [18]).

Proof of the existence of many closed geodesics on M is difficult due to the following: (1) There is a continuous $O(2)$ action on $\Lambda(M)$ that assigns $\sigma \in \Lambda(M)$ and $\alpha \in O(2)$ to the curve $t \rightarrow \sigma(t + \alpha)$, $t \in S^1$; (2) for each integer k and for each critical point $\sigma \in \Lambda(M)$, the curve $t \rightarrow \sigma(kt)$ is a critical point with energy $k^2 E(\sigma)$.

A remarkable result has been obtained by Lyusternik and Shnirel'man [19], which states that there are at least three closed geodesics on every simply connected compact manifold of dimension 2. Fet then proved that there exist at least two closed geodesics on a compact manifold if all critical points of E on $\Lambda(M)$ are nondegenerate [20]. By developing a precise argument concerning the Morse lemma around an isolated degenerate critical point $\sigma \in \Lambda(M)$ of E , Gromoll and Meyer have proved that there exist infinitely many closed geodesics if the sequence of Betti numbers $\{b_i(\Lambda(M))\}$ with respect to any field is unbounded [21]. If $A: M \rightarrow M$ is a certain isometry, then there are also infinitely many A -invariant closed geodesics if the Betti numbers of the space of A -invariant H_1 -curves are not bounded [22, 23]. By investigating the Z_2 -cohomology of $\Lambda(M)$ of compact symmetric spaces, Ziller has proved that if M has the same homotopy type as that of a symmetric space of rank ≥ 2 , then M has infinitely many closed geodesics [24].

Many attempts have been made by W. Klingenberg and others to prove the existence of infinitely many geometrically distinct closed geodesics on every compact Riemannian manifold [25].

References

[1] H. Poincaré, Sur les lignes géodésiques des surfaces convexes, *Trans. Amer. Math. Soc.*, 6 (1905), 237–274. (*Oeuvres*, Gauthier-Villars, 1953, vol. 6, 38–85.)
 [2] G. D. Birkhoff, *Dynamical systems*, Amer. Math. Soc. Colloq. Publ., 1927.
 [3] M. Morse, *Calculus of variations in the large*, Amer. Math. Soc. Colloq. Publ., 1934.

[4] J. Milnor, *Morse theory*, Ann. math. studies 51, Princeton Univ. Press, 1963.
 [5] R. S. Palais, Morse theory on Hilbert manifolds, *Topology*, 2 (1963), 299–340.
 [6] S. S. Chern and R. Lashof, On the total curvature of immersed manifolds, *Amer. J. Math.*, 79 (1959), 306–318.
 [7] A. Phillips, Submersions of open manifolds, *Topology*, 6 (1967), 171–206.
 [8] T. Nagano, *Variational theory in the large* (in Japanese), Kyôritu, 1971.
 [9] M. W. Hirsch, Immersions of manifolds, *Trans. Amer. Math. Soc.*, 93 (1959), 242–276.
 [10] L. Auslander and R. E. MacKenzie, *Introduction to differentiable manifolds*, McGraw-Hill, 1963.
 [11] S. Smale, Generalized Poincaré's conjecture in dimensions greater than four, *Ann. Math.*, (2) 74 (1961), 391–406.
 [12] G. Reeb, Sur certaines propriétés topologiques des variétés feuilletées, *Actualités Sci. Ind.*, 11 (1952), 91–154.
 [13] J. Milnor, *Lectures on the h -cobordism theorem*, Princeton Univ. Press, 1965.
 [14] R. Bott, Nondegenerate critical manifolds, *Ann. Math.*, (2) 60 (1954), 248–261.
 [15] R. Palais and S. Smale, A generalized Morse theory, *Bull. Amer. Math. Soc.*, 70 (1964), 165–172.
 [16] W. Meyer, Kritische Mannigfaltigkeiten in Hilbertmannigfaltigkeiten, *Math. Ann.*, 170 (1967), 45–66.
 [17] J. T. Schwarz, *Nonlinear functional analysis*, Gordon & Breach, 1969.
 [18] L. Lyusternik and A. I. Fet, Variational problems on closed manifolds (in Russian), *Dokl. Akad. Nauk SSSR*, 81 (1951), 17–18.
 [19] L. Lyusternik and L. Shnirel'man, *Méthodes topologiques dans les problèmes variations*, Actualités Sci. Ind., Hermann, 1934.
 [20] A. I. Fet, Variation problems on closed manifolds, *Amer. Math. Soc. Transl.*, ser. 1, 6 (1962), 147–206. (Original in Russian, 1952.)
 [21] D. Gromoll and W. Meyer, Periodic geodesics on compact Riemannian manifolds, *J. Differential Geom.*, 3 (1969), 493–510.
 [22] K. Grove, Condition (C) for the energy integral on certain path spaces and applications to the theory of geodesics, *J. Differential Geom.*, 8 (1973), 207–223.
 [23] M. Tanaka, On the existence of infinitely many isometry-invariant geodesics, *J. Differential Geometry*, 17 (1982) 171–184.
 [24] W. Ziller, Geschlossene Geodätische auf global symmetrischen und homogen Räume, *Bonner Math. Schr.*, 85, 1976.
 [25] W. Klingenberg, *Lectures on closed geodesics*, *Grundlagen der math. Wiss.* 230, Springer, 1978.

280 (XVIII.9) Multivariate Analysis

A. General Remarks

Multivariate analysis consists of methods of statistical analysis of multivariate observations represented by a collection of points in a finite-dimensional Euclidean space \mathbb{R}^p . With the development of powerful computers, multivariate techniques are beginning to be utilized in many fields of science and technology.

B. The Multivariate Linear Model

The multivariate linear model is an immediate extension of the univariate linear model. Suppose that $\mathbf{X} = (\mathbf{X}^{(1)} \dots \mathbf{X}^{(n)})$ denotes the $p \times n$ matrix of n observations of p -dimensional data. Suppose that it can be expressed as

$$\mathbf{X} = \mathbf{B}\mathbf{Z} + \mathbf{U}, \quad (1)$$

where \mathbf{B} is a $p \times m$ matrix of unknown parameters, \mathbf{Z} is a known $m \times n$ matrix of independent variables, and \mathbf{U} is a $p \times n$ matrix of errors. We assume that (i) the \dagger expectations of the elements of \mathbf{U} are zero, that is, the $p \times n$ matrix $E(\mathbf{U}) = 0$, and call the relation (1) a **multivariate linear model**. We usually assume further that (ii) the column vectors $\mathbf{U}^{(i)}$, $i = 1, \dots, n$, of \mathbf{U} are independent and identically distributed, and that (iii) $\mathbf{U}^{(i)}$ is distributed according to a multivariate \dagger normal distribution with \dagger covariance matrix Σ . Analogous to the univariate case, the \dagger least squares estimator $\hat{\mathbf{B}}$ of \mathbf{B} is defined to be the $p \times m$ matrix that minimizes

$$\text{tr}(\mathbf{X} - \mathbf{B}\mathbf{Z})(\mathbf{X} - \mathbf{B}\mathbf{Z})'$$

and is given explicitly by

$$\hat{\mathbf{B}} = \mathbf{X}\mathbf{Z}'(\mathbf{Z}\mathbf{Z}')^{-1} \quad \text{when} \quad |\mathbf{Z}\mathbf{Z}'| \neq 0.$$

Here the symbol $'$ means the transpose of a matrix. Also, an unbiased estimator of Σ is given by

$$\hat{\Sigma} = \mathbf{Q}_e/(n-m), \quad \mathbf{Q}_e = \mathbf{X}\mathbf{X}' - \hat{\mathbf{B}}\mathbf{Z}\mathbf{X}'.$$

$\hat{\mathbf{B}}$ is an unbiased estimator of \mathbf{B} under the assumption (i) above and is the \dagger best linear unbiased estimator under (i) and (ii), while $\hat{\Sigma}$ is unbiased when (i) and (ii) are assumed. Under the assumptions (i)–(iii), $\hat{\mathbf{B}}$ and $\hat{\Sigma}$ form a set of complete \dagger sufficient statistics; hence they are \dagger uniformly minimum variance unbiased estimators. Also under (i)–(iii), elements of $\hat{\mathbf{B}}$ are normally distributed, and their covariance can be expressed by

$$\Sigma \otimes M \quad (\otimes \text{ denotes the } \dagger \text{Kronecker product}),$$

where $M = (\mathbf{Z}\mathbf{Z}')^{-1}$. Applying \dagger Cochran's theorem for the multivariate case, \mathbf{Q}_e is shown to be distributed according to a \dagger Wishart distribution with $n-m$ degrees of freedom.

To test the hypothesis $\mathbf{B} = \mathbf{B}_0$ under (i)–(iii), we put $\mathbf{Q}_B = (\hat{\mathbf{B}} - \mathbf{B}_0)\mathbf{Z}\mathbf{Z}'(\hat{\mathbf{B}} - \mathbf{B}_0)'$ and have $(\mathbf{X} - \mathbf{B}_0\mathbf{Z})(\mathbf{X} - \mathbf{B}_0\mathbf{Z})' = \mathbf{Q}_B + \mathbf{Q}_e$, where \mathbf{Q}_B and \mathbf{Q}_e are independently distributed. The distribution of \mathbf{Q}_B is a Wishart distribution with m degrees of freedom when the hypothesis is true, and a \dagger noncentral Wishart distribution when $\mathbf{B} \neq \mathbf{B}_0$. Based on this fact, several procedures have been proposed. If we require the invariance of procedures with respect to linear transformations of the coordinates of p -dimensional vectors, the roots $\lambda_1, \dots, \lambda_p$ of the \dagger characteristic equation $|\mathbf{Q}_B - \lambda \mathbf{Q}_e| = 0$ form a \dagger maximal invariant statistic; hence the testing procedures should be defined in terms of these roots (\rightarrow 396 Statistic I). Also, the consideration of \dagger power leads to procedures that reject the hypothesis when these roots are large. Commonly used test statistics are (1) the \dagger likelihood ratio test $W = |\mathbf{Q}_e|/|\mathbf{Q}_B + \mathbf{Q}_e| = \prod_i (1 + \lambda_i)^{-1}$ (S. S. Wilks); (2) $\text{tr} \mathbf{Q}_e^{-1} \mathbf{Q}_B = \sum \lambda_i$ (D. N. Lawley and H. Hotelling); (3) $\max \lambda_i$ (S. N. Roy); (4) $\text{tr} \mathbf{Q}_B(\mathbf{Q}_B + \mathbf{Q}_e)^{-1} = \sum \lambda_i(1 + \lambda_i)^{-1}$ (K. C. S. Pillai). Pillai's trace test (4) is locally the most powerful invariant; Wilks's likelihood ratio test (1) has the maximum Bahadur efficiency. The power functions of the tests (1)–(3) have the monotonicity property, namely, they are monotonically nondecreasing with respect to each eigenvalue of $\Omega = \Sigma^{-1}(\mathbf{B} - \mathbf{B}_0)\mathbf{Z}\mathbf{Z}'(\mathbf{B} - \mathbf{B}_0)'$, the matrix of noncentrality parameters for \mathbf{Q}_B . The monotonicity for Pillai's test (4) is known to hold only for restricted cases where the critical value c for the acceptance region $\text{tr} \mathbf{Q}_B(\mathbf{B}_B + \mathbf{Q}_e)^{-1} \leq c$ should satisfy $0 \leq c \leq 1$. All the tests (1)–(4) are unbiased. With respect to 0–1 loss, the tests (1) and (4) are \dagger admissible Bayes and the tests (2) and (3) are \dagger admissible. Tests based on $\min \lambda_i$ are inadmissible.

Small-sample distributions of these statistics are complicated but when $n \rightarrow \infty$, $n \log W$, $n \text{tr} \mathbf{Q}_B \mathbf{Q}_e^{-1}$, and $n \text{tr} \mathbf{Q}_B(\mathbf{Q}_B + \mathbf{Q}_e)^{-1}$ are asymptotically distributed according to a chi-square distribution with pm degrees of freedom under the null hypothesis. Even under the alternative hypothesis that the matrix of noncentrality parameters $\Omega = o(1)$ for large n , the asymptotic distributions remain the same. When $\Omega = O(1)$, they are noncentral chi-square distributions of pm degrees of freedom and noncentrality parameter $\delta = \text{tr} \Omega$. If $\Omega = O(n)$, they are normal distributions, namely, $-\sqrt{n}(\log W + \log |I + \theta|)$, $\sqrt{n}(\text{tr} \mathbf{Q}_B \mathbf{Q}_e^{-1} - \text{tr} \theta)$ and $\sqrt{n}(\text{tr} \mathbf{Q}_B(\mathbf{Q}_B + \mathbf{Q}_e)^{-1} - \text{tr} \theta(I + \theta)^{-1})$ for $\theta = \lim \Omega/n$ have asymptotically normal distributions of zero means and variances given by $2 \text{tr}(I - (I + \theta)^{-2})$, $2 \text{tr}(2\theta + \theta^2)$, and

$2\text{tr}((I+\theta)^{-2}-(I+\theta)^{-4})$, respectively. As a special case, if $m=1$ there exists only one non-zero λ , and the procedures in this paragraph all coincide and are equivalent to one based on $T^2 = M^{-1}(\hat{\mathbf{B}} - B_0)' \hat{\Sigma}^{-1}(\hat{\mathbf{B}} - B_0)$. It is known that under the hypothesis, $(n-p)T^2/(n-1)p$ is distributed according to an F -distribution with $(p, n-p)$ degrees of freedom. When $p=2$ (resp. $m=2$), $(n-m-1)(1-\sqrt{W})/(m\sqrt{W})$ (resp. $(n-p-1)(1-\sqrt{W})/(p\sqrt{W})$) is distributed according to an F -distribution with degrees of freedom $(2m, 2(n-m-1))$ (resp. $(2p, 2(n-p-1))$). Simultaneous \dagger confidence regions of B can be derived from the testing procedures in this paragraph, that is,

$$\text{tr} \mathbf{Q}_e^{-1}(B - \hat{\mathbf{B}})ZZ'(B - \hat{\mathbf{B}})' < c.$$

Moreover, when the matrix B is decomposed as $B = (B_1; B_2)$, where B_1 and B_2 are a $p \times q$ matrix and a $p \times (m-q)$ matrix, respectively, and the hypothesis to be tested is of the form $B_1 = 0$, the test procedures can be obtained as follows: Decompose Z as

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

where Z_1 is a $q \times n$ matrix and Z_2 is an $(m-q) \times n$ matrix, and put

$$\hat{\mathbf{B}}_2^* = \mathbf{X}Z_2'(Z_2Z_2')^{-1},$$

$$\mathbf{Q}^* = \mathbf{X}\mathbf{X}' - \hat{\mathbf{B}}_2^*Z_2'\mathbf{X}', \quad \mathbf{Q}_{B_1} = \mathbf{Q}^* - \mathbf{Q}_e.$$

Then \mathbf{Q}_{B_1} and \mathbf{Q}_e are independent, \mathbf{Q}_{B_1} is distributed according to a Wishart distribution with q degrees of freedom when the hypothesis is true, and we can apply the procedures in the previous paragraph, simply replacing \mathbf{Q}_B by \mathbf{Q}_{B_1} .

Such a procedure is called **multivariate analysis of variance** (or **MANOVA**, for short). Various standard situations can be treated in this way (after some linear transformation of variables, if necessary). Some examples are (1) $\mathbf{X} = (\mathbf{X}^{(1)} \dots \mathbf{X}^{(n)})$, where the $\mathbf{X}^{(i)}$ ($i=1, \dots, n$) are distributed independently according to a p -dimensional normal distribution $N(\boldsymbol{\mu}, \Sigma)$. We can express \mathbf{X} as $\mathbf{X} = \boldsymbol{\mu}\mathbf{1}' + \mathbf{U}$, and the estimators are given by $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \mathbf{X}\mathbf{1}/n$, $\hat{\Sigma} = (\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}')(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}')'/(n-1)$. In this case, we obtain a test for the hypothesis $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ based on **Hotelling's** T^2 statistic, i.e., the test with a \dagger critical region of the form

$$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \hat{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > c.$$

(2) Suppose that $p \times n_i$ matrices \mathbf{X}_i , $i=1, \dots, k$, are samples of size n_i from p -dimensional normal distributions $N(\boldsymbol{\mu}_i, \Sigma)$ with common covariance matrix Σ . The tests for the hypothesis $\boldsymbol{\mu}_1 = \dots = \boldsymbol{\mu}_k$ are obtained from the following observation: Let $\mathbf{Q}_e = \sum_i (\mathbf{X}_i - \bar{\mathbf{X}}_i\mathbf{1}')(\mathbf{X}_i - \bar{\mathbf{X}}_i\mathbf{1})'$, where $\bar{\mathbf{X}}_i = \mathbf{X}_i\mathbf{1}/n_i$, $\mathbf{Q}_\mu = \sum n_i(\bar{\mathbf{X}}_i - \bar{\mathbf{X}})$

$(\bar{\mathbf{X}}_i - \bar{\mathbf{X}})'$, $\bar{\mathbf{X}} = \sum n_i \bar{\mathbf{X}}_i / \sum n_i$. Then $\sum (\mathbf{X}_i - \bar{\mathbf{X}}\mathbf{1}')(\mathbf{X}_i - \bar{\mathbf{X}}\mathbf{1})' = \mathbf{Q}_\mu + \mathbf{Q}_e$. We call \mathbf{Q}_e the **matrix of the sum of squares within classes**, and \mathbf{Q}_μ the **matrix of the sum of squares between classes**.

The latter is distributed according to a Wishart distribution when the hypothesis is true.

(3) Suppose that \mathbf{X}_{ij} are p -dimensional vectors, and that

$$\mathbf{X}_{ij} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + \mathbf{U}_{ij}, \quad \begin{matrix} i=1, \dots, m, \\ j=1, \dots, n, \end{matrix}$$

where $\boldsymbol{\mu}$, $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_j$ are p -dimensional constant vectors such that $\sum \boldsymbol{\alpha}_i = 0$ and $\sum \boldsymbol{\beta}_j = 0$, and the \mathbf{U}_{ij} are independently distributed according to a p -dimensional normal distribution $N(0, \Sigma)$.

We set

$$\mathbf{Q}_\alpha = n \sum (\bar{\mathbf{X}}_{i.} - \bar{\mathbf{X}})(\bar{\mathbf{X}}_{i.} - \bar{\mathbf{X}})',$$

$$\mathbf{Q}_\beta = m \sum (\bar{\mathbf{X}}_{.j} - \bar{\mathbf{X}})(\bar{\mathbf{X}}_{.j} - \bar{\mathbf{X}})',$$

$$\mathbf{Q}_e = \sum \sum (\mathbf{X}_{ij} - \bar{\mathbf{X}}_{i.} - \bar{\mathbf{X}}_{.j} + \bar{\mathbf{X}}) \times (\mathbf{X}_{ij} - \bar{\mathbf{X}}_{i.} - \bar{\mathbf{X}}_{.j} + \bar{\mathbf{X}})',$$

where

$$\bar{\mathbf{X}}_{i.} = \sum_j \mathbf{X}_{ij}/n, \quad \bar{\mathbf{X}}_{.j} = \sum_i \mathbf{X}_{ij}/m,$$

$$\bar{\mathbf{X}} = \sum \sum \mathbf{X}_{ij}/mn.$$

Then we have $\sum \sum (\mathbf{X}_{ij} - \bar{\mathbf{X}})(\mathbf{X}_{ij} - \bar{\mathbf{X}})' = \mathbf{Q}_\alpha + \mathbf{Q}_\beta + \mathbf{Q}_e$, and \mathbf{Q}_α , \mathbf{Q}_β , \mathbf{Q}_e are distributed independently according to (noncentral) Wishart distributions with degrees of freedom $m-1$, $n-1$, and $(n-1)(m-1)$, respectively. The tests for the hypothesis $\boldsymbol{\alpha}_i = 0$ ($i=1, \dots, m$) or $\boldsymbol{\beta}_j = 0$ ($j=1, \dots, n$) are obtained from these matrices.

C. Tests for Covariance Matrices

Let \mathbf{X}_{ij} ($p \times 1$), $j=1, \dots, N_i$, be a random sample from p -variate normal distribution $N(\boldsymbol{\mu}_i, \Sigma_i)$ for $i=1, \dots, k$. For testing the hypothesis $H_0: \Sigma_1 = \dots = \Sigma_k$ with unknown mean vectors $\boldsymbol{\mu}_i$ against all alternatives, the likelihood ratio statistic is given by

$$\frac{N^{N/2} \prod_i |\mathbf{S}_i|^{N_i/2}}{\prod_i N_i^{N_i/2} |\sum \mathbf{S}_i|^{N/2}}, \quad N = N_1 + \dots + N_k, \quad (2)$$

where $\mathbf{S}_i = \sum_{j=1}^{N_i} (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)'$ for $\bar{\mathbf{X}}_i = \sum_{j=1}^{N_i} \mathbf{X}_{ij}/N_i$. If we replace the sample size N_i by the degrees of freedom $n_i = N_i - 1$ and N by $n = N - k$ in (2), the modified likelihood ratio test is unbiased for general p and k [16]. For $k=1$, the hypothesis specifies $H_0: \Sigma_1 = \Sigma_0$ (a given positive definite matrix) and the likelihood ratio statistic is given by $|S|^{N/2} \text{etr}(-\Sigma_0^{-1} S/2)$ ($\text{etr}(x) = \exp(\text{tr } x)$). Again replacing N by $n = N - 1$ yields an unbiased test. Moreover the power function of this modified likelihood ratio test depends only on the eigenvalues of $\Sigma_1 \Sigma_0^{-1}$ and is increasing with respect to the absolute deviation of each

eigenvalue from 1, that is, $|ch_i \Sigma_1 \Sigma_0^{-1} - 1|$. For $k=2$, it is conjectured that the modified likelihood ratio test has a power function monotonically increasing with respect to $|ch_i \Sigma_1 \Sigma_2^{-1} - 1|$. This conjecture has not yet been proved; we know only that the power is increasing if the maximum eigenvalue $ch_1 \Sigma_1 \Sigma_2^{-1}$ increases from 1 or the minimum eigenvalue $ch_p \Sigma_1 \Sigma_2^{-1}$ decreases from 1.

D. Estimation of Mean Vector and Covariance Matrix

Let $\mathbf{X}(p \times 1)$ have p -variate normal distribution $N(\boldsymbol{\mu}, I)$. For estimating the unknown mean, take the sum of squared errors S as a loss $L(\boldsymbol{\mu}, d) = \|\boldsymbol{\mu} - d\|^2$. Since each component X_i of \mathbf{X} is distributed independently according to $N(\mu_i, 1)$ for $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$, it is natural to suppose that \mathbf{X} is a good estimator. In fact \mathbf{X} is a minimax estimator for $\boldsymbol{\mu}$. However, Stein showed that \mathbf{X} is admissible for $p=1$ and 2 but inadmissible for $p \geq 3$. James and Stein showed that the estimator

$$\left(1 - \frac{p-2}{\|\mathbf{X}\|^2}\right)\mathbf{X} \quad (3)$$

dominates \mathbf{X} when $p \geq 3$. This estimator is further dominated by Stein's positive part estimator $(1 - (p-2)/\|\mathbf{X}\|^2)_+ \mathbf{X}$, where $(a)_+$ means a if a is nonnegative and zero if a is negative. The class of estimators $\mathbf{X} + \nabla \log f(\mathbf{X})$, where $\nabla = (\partial/\partial x_1, \dots, \partial/\partial x_p)'$ for $\mathbf{X}' = (x_1, \dots, x_p)$ are all minimax for $\boldsymbol{\mu}$ dominating \mathbf{X} , if $f(\mathbf{X}) \geq 0$ and $\sqrt{f(\mathbf{X})}$ is superharmonic ($\nabla^2 \sqrt{f(\mathbf{X})} \leq 0$), satisfying $E[\|\nabla \log f(\mathbf{X})\|^2] < \infty$ and $E[\Sigma(\partial^2 f(\mathbf{X})/\partial x_i^2)/f(\mathbf{X})] < \infty$. Putting $f(\mathbf{X}) = \|\mathbf{X}\|^{-(p-2)}$ yields the James-Stein estimator (3). For this problem a class of monotone estimators is essentially complete, where an estimator $d(\mathbf{X})$ is called monotone if $d(\mathbf{X}) \leq d(\mathbf{Y})$ whenever $\mathbf{X} \leq \mathbf{Y}$ (defined componentwise). Stein's positive-part estimator is not monotone and is still inadmissible [8].

Let $\mathbf{X}_i(p \times 1)$, $i=1, \dots, n$, be a random sample taken from the p -variate normal distribution $N(\mathbf{0}, \Sigma)$. Then the maximum likelihood estimator for Σ is given by S/n , where $S = \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i'$. S has Wishart distribution $W_p(n, \Sigma)$. To study the estimation problem of Σ , take two loss functions $L_1(\Sigma, d) = \text{tr } d \Sigma^{-1} - \log |d \Sigma^{-1}| - p$ and $L_2(\Sigma, d) = (1/2) \text{tr}(d \Sigma^{-1} - I)^2$. Under L_1 -loss, the best scalar multiple of S is given by S/n , and under L_2 -loss it is given by $S/(n+p+1)$. The James-Stein minimax estimator for L_1 -loss is given by $d_1(S) = K \Delta K'$, where K is the lower triangular matrix with positive diagonal element for which $S = K K'$ and $\Delta = \text{diag}(\Delta_1, \dots, \Delta_p)$ with $\Delta_i = 1/(n+p+1-2i)$. A minimax estimator for

L_2 -loss is given by $d_2(S) = K \Delta K'$ with the same lower triangular matrix K , but Δ is the solution of the linear equation $A \Delta = b$, where A is given by

$$A = \begin{bmatrix} (n+p-1)(n+p+1) & n+p-3 & & & \\ n+p+3 & (n+p-3)(n+p-1) & & & \\ \dots & \dots & \dots & \dots & \\ n-p+1 & n-p+1 & & & \\ & \dots & n-p+1 & & \\ & \dots & n-p+1 & & \\ & \dots & \dots & \dots & \\ & \dots & (n-p+1)(n-p+3) & & \end{bmatrix}$$

and $b' = (n+p-1, n+p-3, \dots, n-p+1)$. The minimax estimators $d_1(S)$ and $d_2(S)$ dominate the best scalar multiples of S . However, they are inadmissible. Let P be a permutation matrix, then the estimator $\Sigma_p P' d_i(PSP')P/p!$ dominates $d_i(S)$ because of the convexity of the loss function L_i for $i=1$ and 2. The estimator $h_1(S) = (S + ub(u)C)/n$, where $u = 1/\text{tr } CS^{-1}$ and C is a fixed positive-definite matrix, dominates S/n under L_1 -loss if $0 \leq b(u) \leq 2(p-1)/n$ and $b(\cdot)$ is nonincreasing. Under L_2 -loss $(S + bC/\text{tr } CS^{-1})/(n+p+1)$ dominates $S/(n+p+1)$, if $0 \leq b \leq 2(p-1)/(n-p+3)$ [7]. The Haff estimator $h_1(S)$ is dominated by the James-Stein estimator $d_1(S)$ if $p \geq 6$. The estimators $(S + b(\text{tr } CS^{-1}/\text{tr } CS^{-2})C)/n$ for $0 \leq b \leq 2(p-1)/n$ under L_1 -loss also dominate S/n and are not dominated by $d_1(S)$. Similar results hold under L_2 -loss.

E. Correlation among Variables

In order to represent the interrelationships among the p variates, the population correlation matrix $P = \{\rho_{ij}\}$ is used. Also, we can calculate the population \dagger multiple correlation coefficient of the i th component X_i and all the rest of the variables by

$$\rho_{i:1\dots(i)\dots p} = \sqrt{1 - P/P_{ii}},$$

and the \dagger partial correlation coefficient of X_i and X_j , given all others, by

$$\rho_{ij:1\dots(i)\dots(j)\dots p} = -P_{ij}/\sqrt{P_{ii}P_{jj}},$$

where P_{ij} are the cofactors of P (\rightarrow 397 Statistical Data Analysis).

The sample correlation matrix is calculated from the data, and the **sample multiple correlation** and the **sample partial correlation coefficients** are calculated from the sample correlation matrix in exactly the same way as the population coefficients are calculated from the population correlation matrix. When $\mathbf{X} = (\mathbf{X}^{(1)} \dots \mathbf{X}^{(n)})$ is a sample of size n from a multivariate normal population, the sampling distributions of $R_{i:1\dots(i)\dots p}$ and $R_{ij:1\dots(i)\dots(j)\dots p}$ are known (\rightarrow 374 Sampling Distributions).

The determinant of the covariance matrix $|\Sigma|$ (or $|\mathbf{S}|$), called the **(sample) generalized variance**, is a measure of the dispersion of a p -dimensional distribution. The distance of two distributions with mean vectors μ_1 and μ_2 , respectively, and with common variance Σ is often expressed by

$$\delta = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2),$$

which is called the **Mahalanobis generalized distance**.

When the data consists of $(p+q)$ -dimensional vectors $\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$ with $q < p$, the interrelation of \mathbf{X} and \mathbf{Y} as a whole can be expressed in the following way: Let the covariance matrix be partitioned as

$$\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{YY} \end{pmatrix}$$

and the nonzero roots of the equation $|\rho \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}| = 0$ be ρ_1, \dots, ρ_q . Then $\rho_1^{1/2}, \dots, \rho_q^{1/2}$, called the **canonical correlation coefficients**, are the maximal invariant parameters with respect to linear transformation of \mathbf{X} and \mathbf{Y} . Also, if we denote the eigenvector corresponding to a root ρ_i by η_i , i.e.,

$$(\rho_i \Sigma_{YY}) \eta_i = (\Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}) \eta_i,$$

the linear function $\eta_i' \mathbf{Y}$ and $\eta_i' \Sigma_{YX} \Sigma_{XX}^{-1} \mathbf{X}$ are called the **canonical variates**.

F. Principal Components

An important problem in multivariate analysis is to express the variations of many variables by a small number of indices. **Principal component analysis** is a technique of dealing with this problem.

Let \mathbf{X} be a $p \times n$ matrix with n column vectors of p -dimensional data. A linear transform of \mathbf{X} , $\mathbf{T} = \mathbf{A}\mathbf{X}$ (\mathbf{A} an $r \times p$ matrix, $r < p$, \mathbf{T} an $r \times n$ matrix) is called the **principal component** if \mathbf{A} is chosen so as to maximize the sum of the squares of the sample multiple correlation coefficients of each of the row vectors of \mathbf{X} to those of \mathbf{T} , namely, if \mathbf{A} is an $r \times p$ matrix formed by the r eigenvectors of the sample correlation matrix of \mathbf{X} corresponding to the r largest eigenvalues or any nonsingular linear transformation of them. This is a characterization of principal components in terms of correlation optimality.

The principal component $\mathbf{T} = \mathbf{A}\mathbf{X}$ is also characterized by the information-loss optimality in that all eigenvalues of $(\mathbf{X} - \mathbf{C}\mathbf{Y} - \mathbf{b}\mathbf{1}_n)'(\mathbf{X} - \mathbf{C}\mathbf{Y} - \mathbf{b}\mathbf{1}_n)$ are simultaneously minimized subject to the condition: \mathbf{C} is a $p \times r$ matrix, \mathbf{b} is a $p \times 1$ vector and \mathbf{Y} is an $r \times n$ matrix. The solution is given by $\mathbf{C}\mathbf{Y} + \mathbf{b}\mathbf{1}_n' = \mathbf{A}'\mathbf{T} + \bar{\mathbf{x}}\mathbf{1}_n'$,

where $\bar{\mathbf{x}} = \mathbf{X}\mathbf{1}_n/n$. The variation optimality of the principal component is given by maximizing simultaneously all the eigenvalues of the matrix $\mathbf{C}'(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}_n')(\mathbf{X} - \bar{\mathbf{x}}\mathbf{1}_n')\mathbf{C}$, subject to the condition that \mathbf{C} is a $p \times r$ matrix such that $\mathbf{C}'\mathbf{C} = \mathbf{I}_r$. The solution is given by $\mathbf{C} = \mathbf{A}'$, namely, $\mathbf{C}'\mathbf{X} = \mathbf{T}$ [15]. If all the correlations between the components of \mathbf{X} are positive, the largest eigenvalue of the covariance matrix of \mathbf{X} is simple and positive. All the coefficients of the first principal component (components of the eigenvector) can be taken to be positive (Perron-Frobenius theorem).

When we assume normality, the eigenvalues of the sample correlation matrix \mathbf{R} are the maximum likelihood estimators of the eigenvalues of the population correlation matrix, and their sampling distributions can be obtained. A hypothesis relevant to principal component analysis is, for example, that the smallest $p-r$ eigenvalues of the correlation matrix are equal, which can be tested by the statistic

$$R_{p-r} = |\mathbf{R}| / (\lambda_1 \dots \lambda_r ((p - \lambda_1 - \dots - \lambda_r) / (p - r))^{p-r}),$$

where $\lambda_1, \dots, \lambda_r$ are the r largest eigenvalues of \mathbf{R} . Under the hypothesis, $-\log R_{p-r}$ (a constant) is asymptotically distributed according to a chi-square distribution when $n \rightarrow \infty$.

Variations of principal component analysis can be obtained by taking the eigenvectors of the covariance matrix of the raw data or of a multiple of it by some weight matrix.

G. Factor Analysis

Factor analysis is closely related to principal component analysis. We assume a model

$$\mathbf{X} = \mathbf{B}\mathbf{F} + \mathbf{U},$$

where \mathbf{B} and \mathbf{F} are unknown $p \times r$ and $r \times n$ matrices of constants ($p > r$) and \mathbf{U} is a $p \times n$ matrix of independent errors. \mathbf{F} is called the matrix of **factor scores** and \mathbf{B} the matrix of **factor loadings**. We assume $\mathbf{F}\mathbf{F}' = n\mathbf{I}$. If $\mathbf{E}(\mathbf{U}\mathbf{U}') = n\Phi$ is known, then by applying the least squares principle, we can determine \mathbf{B} and \mathbf{F} so as to minimize the trace of $(\mathbf{X} - \mathbf{B}\mathbf{F})\Phi^{-1}(\mathbf{X} - \mathbf{B}\mathbf{F})$. Then \mathbf{B} is obtained by taking the r eigenvectors of $\Phi^{-1}\mathbf{X}\mathbf{X}'$ corresponding to the r largest eigenvalues. When Φ is diagonal but unknown, we can solve the simultaneous equation for \mathbf{B} and Φ , whose solutions are the matrix $\hat{\mathbf{B}}$ with columns equal to eigenvectors of $\hat{\Phi}^{-1}\mathbf{X}\mathbf{X}'$ and the diagonal matrix $\hat{\Phi}$ with elements equal to the diagonal elements of $\mathbf{X}\mathbf{X}'/n - \hat{\mathbf{B}}\hat{\mathbf{B}}'$.

If we assume that \mathbf{U} is normal, the procedure in the preceding paragraph for the case

when Φ is known is equivalent to the maximum likelihood method. When Φ is unknown, we can assume further that the columns in F are also multivariate normal vectors distributed independently of U , which implies that the columns of X are also normal vectors with the covariance matrix $\Sigma = BB' + \Phi$. B and Φ are estimated from the sample covariance matrix, and the solutions of the simultaneous equation for B and Φ gives the maximum likelihood estimators. However, there is the so-called **identification** problem of determining whether for given Σ and r the decomposition $\Sigma = BB' + \Phi$ is unique. This problem has not yet been completely settled. If the solution $\hat{\Phi}$ is positive definite and $\hat{B}\hat{B}'$ is positive semidefinite, it is called proper. When Φ is not positive definite, it is called a Haywood case, and when $\hat{B}\hat{B}'$ is not positive semidefinite it is called a complex case. Sometimes iterative procedures lead to improper solutions.

H. Canonical Correlation Analysis

Canonical correlation analysis can also be used for descriptive purposes. Sample canonical variates have various descriptive implications. Suppose that $\eta'_1 Y$ and $\xi'_1 X$ are the first canonical variates corresponding to the largest canonical correlation $\rho_1^{1/2}$. Then $\rho_1^{1/2}$ is the largest possible correlation between a linear function of X and a linear function of Y and is actually equal to the correlation of $\eta'_1 Y$ and $\xi'_1 X$. Similarly, the second canonical correlation is equal to the largest possible correlation between linear functions in X and in Y which are orthogonal to $\xi'_1 X$ and $\eta'_1 Y$, respectively, and so forth.

As a second interpretation of the canonical variates, we consider the linear regression model

$$Y = BX + U,$$

where Y , B , U are $q \times n$, $q \times p$, $q \times n$ matrices, respectively, and rank $B = r < q$. Then there exist an $r \times p$ matrix A and a $q \times r$ matrix C satisfying $C'\Sigma^{-1}C = I$ such that $B = CA$, where $E(UU') = n\Sigma$. Putting $T = AX$, we get a regression model $Y = CT + U$, regarding T as a matrix of regressor variables. If Σ is known, least squares considerations lead to minimizing $\text{tr}(Y - BX)' \Sigma^{-1} (Y - BX)$ with the condition rank $B = r$. The resulting row vectors of A consist of the r eigenvectors corresponding to the r largest eigenvalues of the matrix $S_{xx}^{-1} S_{xy} \Sigma^{-1} S_{yx}$ and column vectors of C consist of the r eigenvectors corresponding to the r largest eigenvalues of the matrix $S_{yx} S_{xx}^{-1} S_{xy} \Sigma^{-1}$. If Σ is replaced by S_{yy} , $T = AX$ and $Z = C'S_{yy}^{-1} Y$ are equal to the matrices of canonical variates.

If we assume that U is normal, this procedure is equivalent to the maximum likelihood method. It should be remarked that although the model here is not symmetric in X and Y , the results are symmetric in X and Y , and therefore they will be the same if the roles of X and Y are interchanged in this model.

I. Linear Discriminants and Problems of Classification

Let $p \times n_i$ matrices X_i ($i = 1, \dots, k$) be the set of observations for k distinct populations with a common covariance matrix. We determine a vector a such that $T_i = a'X_i$ reveals the differences of the k populations as much as possible, or, more precisely, so that the ratio of the sum of squares between classes of T to the sum of squares within classes is maximized. If the matrices of the sums of squares between and within classes are Q_b and Q_w , respectively, the ratio is equal to

$$l = a'Q_b a / a'Q_w a,$$

which is maximized when a is equal to the eigenvector of $Q_w^{-1}Q_b$ corresponding to the largest eigenvalue. The linear function $t = a'X$ is called the **linear discriminant function**. When $k = 2$, a is given by $a = Q_w^{-1}(\bar{X}_1 - \bar{X}_2)$, where \bar{X}_1 and \bar{X}_2 are sample mean vectors. When $k > 2$, we let A be the matrix formed by the r eigenvectors corresponding to the first r largest eigenvalues of $Q_w^{-1}Q_b$, and set $T_i = AX_i$. From this we can construct the r -dimensional discriminant function. These functions can be used to locate the k populations in r -dimensional space, and also to decide to which population a new observation belongs. For the latter problem we can also construct k quadratic functions $s_i = (X - \bar{X}_i)'Q_w^{-1}(X - \bar{X}_i)$, where \bar{X}_i is the sample mean vector of the i th population, and X a new observation. Then we can decide whether X belongs to the population corresponding to the minimum s_i . Such a method is called a **classification procedure**.

J. Discrete Multivariate Analysis

Let X_{ijk} be an observed frequency on three characteristics, each belonging to ijk th class for $1 \leq i \leq I$, $1 \leq j \leq J$ and $1 \leq k \leq K$. Assume that X_{ijk} is a sample from multinomial distribution having p_{ijk} as a probability of occurrence for an (i, j, k) cell, where $p_{...} = \sum p_{ijk} = 1$. The multinomial observation X_{ijk} with probability p_{ijk} is called an $I \times J \times K$ contingency table. If we further assume that

$$\log p_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha_{ij} + \beta_{jk} + \gamma_{ki} + \delta_{ijk}$$

with the restrictions on parameters $\sum_i \alpha_i = \sum_i \alpha_{ij} = \sum_j \alpha_{ij} = 0$, similarly on β 's and γ 's and $\sum_i \delta_{ijk} = \sum_j \delta_{ijk} = \sum_k \delta_{ijk} = 0$, we say that a saturated log-linear model is given. Here the number of observations and the number of parameters are equal, and no errors can be estimated. The parameter δ_{ijk} is called the three-factor effect, or equivalently, the second-order interaction. Similarly, α_{ij} , β_{jk} , γ_{ki} are called the two-factor effects or the first-order interactions. Finally α_i , β_j , γ_k are called the main effects. A simple model is obtained when all the first- and second-order interactions vanish. This is equivalent to the independence model: $p_{ijk} = p_{i..} p_{.j.} p_{.k.}$ and the maximum likelihood estimators for μ , α_i , β_j , γ_k are obtained from $\hat{p}_{ijk} = \mathbf{X}_{i.} \mathbf{X}_{.j.} \mathbf{X}_{.k.} / n^3$, where $n = \mathbf{X}_{...}$. A nontrivial model is obtained by putting all the second-order interactions equal to zero. The likelihood equations for this model are given by

$$np_{ij.} = \mathbf{X}_{ij.}, \quad np_{i.k} = \mathbf{X}_{i.k}, \quad np_{.jk} = \mathbf{X}_{.jk}. \tag{4}$$

Bartlett first described a solution of (4) when $I = J = K = 2$. For a $2 \times 2 \times 2$ table, the solution can be expressed by $\hat{p}_{ijk} = (x_{ijk} \pm \theta) / n$ because of the constraints for $\mathbf{X}_{ij.}$, $\mathbf{X}_{i.k}$, and $\mathbf{X}_{.jk}$. Putting $\delta_{111} = \theta$ for \hat{p}_{ijk} yields a cubic equation for θ :

$$\begin{aligned} &(\mathbf{X}_{111} + \theta)(\mathbf{X}_{122} + \theta)(\mathbf{X}_{221} + \theta)(\mathbf{X}_{212} + \theta) \\ &= (\mathbf{X}_{121} - \theta)(\mathbf{X}_{112} - \theta)(\mathbf{X}_{211} - \theta)(\mathbf{X}_{222} - \theta). \end{aligned}$$

For a general $I \times J \times K$ table, the equations (4) have a unique solution within the no-three-factor effect model if there exists $q_{ijk} > 0$ satisfying (4). The unique solution maximizes the likelihood. To solve the likelihood equations (4), standard iterative procedures, such as the Newton-Raphson method, can be applied. However, the following iterative scaling method of Deming and Stephan is more useful:

$$\begin{aligned} np_{ijk}^{(3m+1)} &= p_{ijk}^{(3m)} \frac{\mathbf{X}_{ij.}}{p_{ij.}^{(3m)}}, \\ np_{ijk}^{(3m+2)} &= p_{ijk}^{(3m+1)} \frac{\mathbf{X}_{i.k}}{p_{i.k}^{(3m+1)}}, \\ np_{ijk}^{(3m+3)} &= p_{ijk}^{(3m+2)} \frac{\mathbf{X}_{.jk}}{p_{.jk}^{(3m+2)}}. \end{aligned} \tag{5}$$

The first iteration adjusts $p_{ijk}^{(3m)}$ by fitting proportionally with respect to k so that $np_{ij.}^{(3m+1)}$ is equal to $\mathbf{X}_{ij.}$, and similarly for the second and third iterations. Starting with any initial values satisfying the first-order interaction model, the iterative scaling method (5) converges to the unique solution of (4) as $m \rightarrow \infty$.

K. Other Problems

The sampling distributions associated with the procedures discussed here are usually very

complicated, and often only asymptotic properties are known (\rightarrow 374 Sampling Distributions). Nonparametric rank analogs of many multivariate techniques for normal distribution are found in [18]. Robustness of the distributions of test statistics or of latent roots is investigated in [3, 10, 14]. Multivariate data analysis is found in [5].

References

[1] T. W. Anderson, An introduction to multivariate statistical analysis, Wiley, second edition, 1984.
[2] Y. M. M. Bishop, S. E. Fienberg, and P. W. Holland, Discrete multivariate analysis, MIT Press, 1976.
[3] Y. Fujikoshi, Asymptotic expansions for the distributions of the sample roots under nonnormality, *Biometrika*, 67 (1980), 45–51.
[4] N. C. Giri, Multivariate statistical inference, Academic Press, 1977.
[5] R. Gnanadesikan, Methods for statistical data analysis of multivariate observations, Wiley, 1977.
[6] S. J. Haberman, The analysis of frequency data, Univ. of Chicago Press, 1974.
[7] L. R. Haff, Empirical Bayes estimation of the multivariate normal covariance matrix, *Ann. Statist.*, 8 (1980), 586–597.
[8] I. Hashimoto, Estimation for distributions with monotone likelihood ratio: Case of vector-valued parameter, *Commun. Statist.*, A10 (1981), 1901–1913.
[9] W. James and C. Stein, Estimation with quadratic loss, *Proc. Fourth Berkeley Symp. Math. Statist. and Prob.*, 1 (1961), 361–379.
[10] T. Kariya, Robustness of multivariate tests, *Ann. Statist.*, 9 (1981), 1267–1275.
[11] M. G. Kendall, Multivariate analysis, Griffin, 1975.
[12] A. M. Kshirsagar, Multivariate analysis, Dekker, 1972.
[13] D. N. Lawley and A. E. Maxwell, Factor analysis as a statistical method, Butterworth, 1963.
[14] R. J. Muirhead and C. M. Waternaux, Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations, *Biometrika*, 67 (1980), 31–43.
[15] M. Okamoto, Optimality of principal components, *Multivariate Analysis II*, P. R. Krishnaiah (ed.), Academic Press, 1968, 673–685.
[16] M. D. Perlman, Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate normal populations, *Ann. Statist.*, 8 (1980), 247–263.

- [17] M. D. Perlman and I. Olkin, Unbiasedness of invariant tests for MANOVA and other multivariate problems, *Ann. Statist.*, 8 (1980), 1326–1341.
- [18] M. L. Puri and P. K. Sen, *Nonparametric methods in multivariate analysis*, Wiley, 1971.
- [19] C. R. Rao, *Linear statistical inference and its applications*, Wiley, 1973.
- [20] R. Schwartz, Admissible tests in multivariate analysis of variance, *Ann. Math. Statist.*, 38 (1967), 698–710.
- [21] M. S. Srivastava and C. G. Khatri, *An introduction to multivariate statistics*, North-Holland, 1979.
- [22] C. Stein, Estimation of the mean of a multivariate normal distribution, *Ann. Statist.*, 9 (1981), 1135–1151.

N

281 (XIX.5) Network Flow Problems

A. Introduction

The **network flow problem** is a special kind of mathematical-programming problem (\rightarrow 255 Linear Programming, 264 Mathematical Programming, 292 Nonlinear Programming), where the variables of the objective function and the constraints are all defined in terms of a graph (\rightarrow 186 Graph Theory). Owing to its special structure, the mathematical properties of the network flow problem as well as the solution algorithms have been investigated in detail. Network flow problems have a variety of useful applications in fields such as transportation, scheduling, and resource allocation, and in operations research in general, so that they now constitute an important class of mathematical programming. By the term "network" (*Netzwerk* in German, *réseau* in French, *сет'* in Russian) we usually mean a graph with some physical attributes attached to its edges and vertices.

B. Basic Form of the Problem

Let $G = (V, E, \delta^+, \delta^-)$ be a graph with vertex set V , edge set E , and incidence relations δ^+ , $\delta^-: E \rightarrow V$, and let \mathbf{R} be the field of real numbers. (Most of the statements in the following are valid if we take for \mathbf{R} an ordered field or ordered additive group more general than the field of real numbers.) Furthermore, we regard the collection Ξ of all functions $\xi: E \rightarrow \mathbf{R}$ as a vector space of dimension $|E|$, denoting $\xi(e_\kappa)$ by ξ^κ if $E = \{e_1, \dots, e_n\}$. Similarly, the collection Ω of all functions $\omega: V \rightarrow \mathbf{R}$ is a vector space of dimension $|V|$. If we define the mapping $\delta^\pm: V \rightarrow 2^E$ by $\delta^\pm v = \{e \mid \delta^\pm e = v\}$ (\rightarrow 186 Graph Theory), a linear mapping $\delta: \Xi \rightarrow \Omega$ is naturally introduced through the relation $(\delta\xi)(v) = \sum_{e \in \delta^+ v} \xi(e) - \sum_{e \in \delta^- v} \xi(e)$. A vector $\xi \in \Xi$ is called a **flow** on G if $\delta\xi = 0$. The dual space Ξ^* of Ξ and the dual Ω^* of Ω are identified with the collection of functions $\eta: E \rightarrow \mathbf{R}$ and that of $\zeta: V \rightarrow \mathbf{R}$, respectively, under the obvious correspondence. Then the linear mapping $\delta: \Omega^* \rightarrow \Xi^*$, which is contragredient to δ , is defined by means of the relation $(\delta\zeta)(e) = \zeta(\delta^+ e) - \zeta(\delta^- e)$. A vector $\eta \in \Xi^*$, which is the image of ζ under δ , i.e., which satisfies $\eta = \delta\zeta$, is called a **tension** on G , and ζ is called the **potential** on G corresponding to the tension η . (Sometimes, $\xi(e)$ is called a flow **in** edge e , $\eta(e)$ a tension **across** e , and $\zeta(v)$ a potential **at** vertex v .)

A continuous curve $C (\subset \mathbf{R} \times \mathbf{R})$ on the

Euclidean plane $\mathbf{R} \times \mathbf{R}$ is said to be **monotone** if $(x_1 - x_2)(y_1 - y_2) \geq 0$ for any $(x_1, y_1), (x_2, y_2) \in C$. A monotone curve C is called a **characteristic curve** if its projection to each coordinate axis, i.e., $C_x = \{x \mid (x, y) \in C\}$ and $C_y = \{y \mid (x, y) \in C\}$, is a closed interval. For a characteristic curve C , two convex functions, $\varphi(x) = \int_{x_0}^x (y - y_0) dx$ and $\psi(y) = \int_{y_0}^y (x - x_0) dy$, are defined, where (x_0, y_0) is a point fixed on C and the integrals are taken along $(x, y) \in C$. (It is understood that $\varphi(x) = \infty$ if $x \notin C_x$ and $\psi(y) = \infty$ if $y \notin C_y$.) These two functions defined for a fixed C are conjugate to each other in Fenchel's sense, and they satisfy $\varphi(x) + \psi(y) \geq (x - x_0)(y - y_0)$ for any $(x, y) \in \mathbf{R} \times \mathbf{R}$, where the inequality reduced to an equality if and only if $(x, y) \in C$ (\rightarrow 88 Convex Analysis, 292 Nonlinear Programming).

A network N in network flow theory is a graph G to each edge e_κ whose edge set $E = \{e_1, \dots, e_n\}$ is given a characteristic curve C^κ . On a network N , the following three problems, PI, PII, and PIII, are defined. PI: Find a flow ξ that minimizes $\Phi(\xi) \equiv \sum_{\kappa=1}^n \varphi_\kappa(\xi^\kappa)$ under the constraints $\xi^\kappa \equiv \xi(e_\kappa) \in C_x^\kappa$ ($\kappa = 1, \dots, n$). PII: Find a tension η which minimizes $\Psi(\eta) \equiv \sum_{\kappa=1}^n \varphi^\kappa(\eta_\kappa)$ under the constraints $\eta_\kappa \equiv \eta(e_\kappa) \in C_y^\kappa$ ($\kappa = 1, \dots, n$). PIII: Find a pair (ξ, η) of a flow ξ and a tension η such that $(\xi^\kappa, \eta_\kappa) \in C^\kappa$ for all $\kappa = 1, \dots, n$. (With respect to PI or PII, a flow or a tension satisfying the "constraints" is ordinarily called a **feasible** flow or tension, respectively.) Then, as a special case of the (Karush-)Kuhn-Tucker theorem and the duality theorem in nonlinear programming, we have the following theorems (A) and (B).

(A) For a given network N , one and only one of the following four alternatives is the case:

- (i) There is a feasible flow in PI, and at the same time, there is a feasible tension in PII. In this case, both PI and PII have a solution, and, for any solution $\hat{\xi}$ of PI and any solution $\hat{\eta}$ of PII, the pair $(\hat{\xi}, \hat{\eta})$ is a solution of PIII; and, conversely, for any solution $(\hat{\xi}, \hat{\eta})$ of PIII, the flow $\hat{\xi}$ is a solution of PI and the tension $\hat{\eta}$ is a solution of PII.
- (ii) There is a feasible tension in PII, whereas there is no feasible flow in PI. In this case, $\Phi(\xi)$ of PI is not bounded below for feasible tensions ξ .
- (iii) There is a feasible tension in PII, whereas there is no feasible flow in PI. In this case, $\Psi(\eta)$ of PII is not bounded below for feasible tensions η .
- (iv) There is neither a feasible flow in PI nor a feasible tension in PII.

(B) Let $C_x^\kappa = [b^\kappa, c^\kappa]$ ($b^\kappa \leq c^\kappa$; b^κ can be $-\infty$, and c^κ, ∞) and $C_y^\kappa = [d_\kappa, f_\kappa]$ (d_κ can be $-\infty$, and f_κ, ∞). Then a necessary and sufficient condition for PI to have a feasible flow is that

$\sum_1 c^k - \sum_2 b^k \geq 0$ for every cutset (cocycle or cocircuit) ω of G , where the summation \sum_1 (resp. \sum_2) is taken over all the edges e_k ($\in E$) that lie in ω in the positive (resp. negative) direction. Similarly, a necessary and sufficient condition for PII to have a feasible tension is that $\sum_1 f_k - \sum_2 d_k \geq 0$ for every tieset (cycle or circuit) θ of G , where \sum_1 (resp. \sum_2) is taken over all the edges e_k that lie in θ in the positive (resp. negative) direction.

C. Shortest Paths, Maximum Flows, and Minimum-Cost Flows

We choose one of the edges of G , say e_1 , as the **reference edge** and assign it the parametric characteristic curve $C^1 \langle \alpha \rangle = \{(x, y) | y = x + \alpha\}$. Then if there is a feasible flow on the network obtained from G by contracting (i.e., short-circuiting) e_1 , and if at the same time there is a feasible tension on the network obtained from G by deleting (i.e., open-circuiting) e_1 , then problem PIII has a solution $(\xi \langle \alpha \rangle, \eta \langle \alpha \rangle)$ for every real α , and $\xi^1 \langle \alpha \rangle$ and $\eta_1 \langle \alpha \rangle$ are uniquely determined for each α . The problem of determining these parametric solutions is the **two-terminal problem** for the **two-terminal network** N_1 (which is obtained from G by deleting e_1) with the vertex $\partial^- e_1$ as the **entrance** (or **source**) and the vertex $\partial^+ e_1$ as the **exit** (or **sink**). The curve $C = \{(\xi^1 \langle \alpha \rangle, \eta_1 \langle \alpha \rangle) | \alpha \in \mathbb{R}\}$ enjoys the properties of a characteristic curve, and is called the **two-terminal characteristic** of N_1 with respect to the entrance $\partial^- e_1$ and exit $\partial^+ e_1$. A two-terminal network for which only the projections to the x -axis $C_x^k = [b^k, c^k]$ of the edge characteristics are specified ($k = 2, \dots, n$) is called a **capacitated network**, and the **maximum-flow problem** for a capacitated network N_1 can be mathematically formulated as the problem of determining the projection to the x -axis $C_x = [b, c]$ of the two-terminal characteristics of N_1 . For the maximum-flow problem, the relation $c = \min\{\sum_1 c^k - \sum_2 b^k\}$ holds, where the minimum is taken over all the cutsets that contain e_1 in the negative direction and where \sum_1 (resp. \sum_2) denotes the summation over all the edges, except e_1 , lying in a cutset in the positive (resp. negative) direction. This relation is called the **maximum-flow minimum-cut theorem**. (A similar relation holds also for b .)

Similarly, or dually, the problem of determining the y -projection $C_y = [d, f]$ of the two-terminal characteristic of a two-terminal network N_1 for which the y -projections $C_y^k = [d_k, f_k]$ of the characteristics are specified to the nonreference edges e_k ($k = 2, \dots, n$) is a network flow formulation of the **shortest-path problem**. For the shortest-path problem, the

relation $f = \min\{\sum_1 f_k - \sum_2 d_k\}$, called the **maximum-separation minimum-distance theorem**, holds, where the minimum is taken over all the tiesets that contain e_1 in the positive direction and where \sum_1 (resp. \sum_2) denotes the summation over all the edges, except e_1 , lying in a tieset in the positive (resp. negative) direction.

The **minimum-cost flow problem** is to determine the two-terminal characteristic of N_1 when all the C^k ($k \neq 1$) are of staircase form.

A number of algorithms exist for time complexity (\rightarrow 71 Complexity of Computations) $O(|V|^3)$ for the shortest-path problem [4]; the algorithm proposed by E. W. Dijkstra [7] for the case $d_k \leq 0 \leq f_k$ for all k ($k \neq 1$) is of complexity $O(|V|^2)$. The algorithm of A. V. Karzanov [8, 9] for the maximum-flow problem is of complexity $O(|V|^3)$. The minimum-cost flow problem can be solved by alternately solving the subproblems of the shortest-path type and those of the maximum-flow type, but no algorithm of time complexity polynomial in $|E|$ and $|V|$ has been found.

D. Transportation and Scheduling

Let us make the edges of a graph $G = (V, E, \partial^+, \partial^-)$ correspond to the transportation routes and the flow to the stream of a commodity; impose a **capacity constraint** of the form $0 \leq \xi^k \leq c^k$ on the flow ξ^k in each edge e_k , and assume a **cost** function of the form $\varphi_k(\xi^k) = f_k \xi^k$ for each edge. (c^k is a constant called the **capacity** of edge e_k , and f_k is a constant called the **unit cost** of edge e_k .) Furthermore, let us specify a subset V_1 ($\subset V$) of vertices as the set of entrances and another subset V_2 ($\subset V$) as the set of exits, where $V_1 \cap V_2 = \emptyset$, and prescribe the amount of inflow $q(v)$ to each entrance $v \in V_1$ and the amount of outflow $q(v)$ from each exit $v \in V_2$, where we must have $\sum_{v \in V_1} q(v) = \sum_{v \in V_2} q(v)$. Planning a transportation plan that satisfies all the above-prescribed conditions and that minimizes the total cost $\sum_{k \in E} \varphi_k(\xi^k)$ can be reduced to finding a minimum-cost flow on the extended network $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{\partial}^+, \tilde{\partial}^-)$, defined as follows, such that the flow in the reference edge is to be maximum: $\tilde{V} = V \cup \{s\} \cup \{t\}$ ($s, t \notin V$), $\tilde{E} = E \cup E_1 \cup E_2 \cup \{e_0\}$ ($E_1 \cap E = E_2 \cap E = \emptyset$, $e_0 \notin E \cup E_1 \cup E_2$; e_0 is the reference edge, $\tilde{\partial}^+ e_0 = t$, $\tilde{\partial}^- e_0 = s$; $E_1 = \{e | \tilde{\partial}^+ e = s, \tilde{\partial}^- e = v, v \in V_1\}$, $C(e) = \{(0, y) | y \leq 0\} \cup \{(x, 0) | 0 \leq x \leq q(\tilde{\partial}^- e)\} \cup \{(q(v), y) | 0 \leq y\}$, where $\tilde{\partial}^- e = v \in V_1$; $E_2 = \{e | \tilde{\partial}^+ e = v, \tilde{\partial}^- e = t, v \in V_2\}$, $C(e) = \{(0, y) | y \leq 0\} \cup \{(x, 0) | 0 \leq x \leq q(v)\} \cup \{(q(v), y) | 0 \leq y\}$, where $\tilde{\partial}^+ e = v \in V_2$), $\tilde{\partial}^\pm |_E = \partial^\pm$, $C(e_k) = \{(0, y) | y \leq f_k\} \cup \{(x, f_k) | 0 \leq x \leq c^k\} \cup \{(c^k, y) | y \geq f_k\}$ for $e_k \in E$.

For the project-scheduling problem with an

Network Flow Problems

acyclic graph $G = (V, E, \delta^+, \delta^-)$ as the **arrow diagram** for which the **start node** $s \in V$ and the **completion node** are specified, we make an edge $e \in E$ correspond to a **job** (or **activity**), and a vertex correspond to the **event** that all the jobs corresponding to $\delta^- v$ have been finished, whereas those corresponding to $\delta^+ v$ are ready to commence. Then we interpret the negative of the tension of an edge as the time (i.e., duration) spent on the corresponding job and the potential at a vertex as the time (i.e., instant) of the corresponding event taking place. Furthermore, we assume that to each edge e_k (or the corresponding job) are given the **normal duration** $-d_k (> 0)$, the **crash duration** $-f_k (> 0, < -d_k)$, and the unit cost $c^k (> 0)$ we have to pay in order to decrease the job time by one unit, where the job time necessarily lies between the normal and the crash duration. Under the above-listed specifications, we consider the relation between the total project time (i.e., the duration from the event corresponding to the start node to that corresponding to the completion node) and the extra cost to be paid for decreasing the project time below the normal one. This problem can be reduced to the two-terminal characteristic problem for the following network \tilde{G} : \tilde{G} is obtained from G by adding the reference edge e_0 ($\tilde{\delta}^+ e_0 = t, \tilde{\delta}^- e_0 = s$); $C^* = \{(0, y) | y \leq d_k\} \cup \{(x, d_k) | 0 \leq x \leq c^k\} \cup \{(c^k, y) | d_k \leq y \leq f_k\} \cup \{(x, f_k) | c^k \leq x\}$.

E. Applications to Combinatorial Optimization

Many problems in combinatorial optimization can be reduced to network flow problems. The problem of finding a maximum matching on a bipartite graph G (\rightarrow 186 Graph Theory) is reduced to the maximum-flow problem for the graph representing the transportation problem on G with one of the two vertex sets as the entrance set and the other as the exit set, where all the edges have a unit capacity and the amount of inflow/outflow to/from each entrance/exit vertex is equal to unity (the cost being irrelevant). (The existence of an integer solution to this kind of maximum-flow problem is proved constructively on the basis of the solution algorithm.) The maximum-flow minimum-cut theorem in this particular case can be stated as follows: The maximum cardinality of matchings on a bipartite graph is equal to the minimum cardinality of vertex subsets which cover all the edges (this is known as the **König-Egerváry theorem**). The Dilworth theorem for a partially ordered set, the criterion for the existence of a graph with

prescribed vertex degrees, etc., are known to be reducible to network flow problems [2].

F. Generalizations

The network flow problem may be generalized in various directions. Replacing a graph by a matroid (\rightarrow 66 Combinatorics) or considering stronger conditions on the feasibility of flows and tensions are natural extensions [4, 6, 10, 11]. Another extension is to consider several kinds of flow, instead of a single kind, that simultaneously affect the capacities of edges. This latter problem is called the **multicommodity flow problem** in contrast to the **single-commodity flow problem** that was treated above [5]. It can be said that any extension of the network flow problem aims at a mathematical model that has wider application without losing the advantage of having simple effective solution algorithms.

References

- [1] G. J. Minty, Monotone networks, Proc. Roy. Soc. London, A257 (1960), 194–212.
- [2] L. R. Ford, Jr., and D. R. Fulkerson, Flows in networks, Princeton Univ. Press, 1962.
- [3] C. Berge and A. Ghoulia-Houri, Programmes, jeux et réseaux de transport, Dunod, 1962.
- [4] M. Iri, Network flow, transportation and scheduling—Theory and algorithms, Academic Press, 1969.
- [5] T. C. Hu, Integer programming and network flows, Addison-Wesley, 1969.
- [6] E. L. Lawler, Combinatorial optimization—Networks and matroids, Holt, Rinehart and Winston, 1976.
- [7] E. W. Dijkstra, A note on two problems in connexion with graphs, Numerische Math., 1 (1959), 269–271.
- [8] E. A. Dinits, Algorithm solutions to maximum network flow problems by power analysis (in Russian), Dokl. Akad. Nauk SSSR, 194 (1970), 754–757.
- [9] A. V. Karzanov, Calculation of maximum network flow by power methods (in Russian), Dokl. Akad. Nauk SSSR, 215 (1974), 49–52.
- [10] M. Iri and S. Fujishige, Use of matroids in operations research, circuits and systems theory, Intern. J. Syst. Sci., 12 (1981), 27–54.
- [11] C. P. Bruter, Eléments de la théorie des matroïdes, Lecture notes in math. 387, Springer, 1974.

282 (XX.17) Networks

A. Linear Graphs (\rightarrow 186 Graph Theory)

A **linear graph** (or simply **graph**) is an object composed of (i) a finite set $\{B_\kappa\}$ ($\kappa = 1, \dots, n$) of elements called **branches** (or **edges**), (ii) a finite set $\{N_a\}$ ($a = 1, \dots, m$) of elements called **nodes** (or **vertices**), and (iii) an **incidence relation** between branches and nodes represented by a function $[B_\kappa : N_a]$ from $\{B_\kappa\} \times \{N_a\}$ to $\{0, 1, -1\}$ such that for every κ , there exist exactly one N_a with $[B_\kappa : N_a] = 1$ and exactly one N_a with $[B_\kappa : N_a] = -1$, with all the other $[B_\kappa : N_a]$ equal to 0. (Intuitively, a branch B_κ starts from the node N_a with $[B_\kappa : N_a] = 1$ and ends at the node N_b with $[B_\kappa : N_b] = -1$.) In terms of topology, a (linear) graph is a 1-dimensional finite \dagger simplicial complex (\rightarrow 70 Complexes, 186 Graph Theory). A **network** in the wide sense is a linear graph whose branches and nodes are endowed with some physical properties.

B. Networks

A **contact network**, one of the simplest kinds of networks, is an abstraction of a circuit whose branches correspond to contact points of relays and switches that are allowed to take a finite number of physical states, e.g., the two states "on" and "off." The theory of contact networks is developed by means of \dagger Boolean algebra and is applied to switching networks, such as telephone exchange networks and the logical networks of digital computers.

In most cases, to the branches B_κ of a network two kinds of real quantities i_κ and E_κ are assigned (which may be variables or functions of time) satisfying the conditions:

$$(i) \sum_{\kappa=1}^n [B_\kappa : N_a] i_\kappa = 0 \quad (a = 1, \dots, m);$$

(ii) there exist E_a such that

$$\sum_{a=1}^m [B_\kappa : N_a] E_a = E_\kappa \quad (\kappa = 1, \dots, n).$$

In the case of **electric networks**, where i_κ , E_κ , and E_a are the current in branch B_κ , the voltage across branch B_κ , and the potential at node N_a , respectively, these conditions are known as **Kirchhoff's laws**.

The **network flow problem**, which has a number of practically important applications in \dagger operations research (e.g., transportation problems, project-scheduling problems) and is a special case of \dagger mathematical programming,

can be formulated as the problem of minimizing $\sum_{\kappa=1}^m f_\kappa(i_\kappa)$ under condition (i) (or minimizing $\sum_{\kappa=1}^n f_\kappa(E_\kappa)$ under condition (ii)), where for each branch B_κ , f_κ is a given \dagger convex function defined on a given interval $[a_\kappa, b_\kappa]$.

C. Electric Networks

Since there has been a great deal of research on electric networks, "network" often means "electric network." We call a branch in which the current is a given function of time a current source, and one across which the voltage is a given function of time a voltage source. A branch that is either a current source or a voltage source is called a **source branch**. A network with M source branches is called an **M-port network**. If the currents i_κ in and the voltages E_κ across the non-source branches (n' in number) are related by

$$E_\kappa = \sum_{\lambda=1}^{n'} z_{\kappa\lambda} i_\lambda \quad (\kappa = 1, \dots, n')$$

or

$$i_\kappa = \sum_{\lambda=1}^{n'} y_{\kappa\lambda} E_\lambda \quad (\kappa = 1, \dots, n'),$$

where the $z_{\kappa\lambda}$ or $y_{\kappa\lambda}$ are linear \dagger integrodifferential operators, the network is said to be **linear**. If $z_{\lambda\kappa} = z_{\kappa\lambda}$ or $y_{\kappa\lambda} = y_{\lambda\kappa}$, it is said to be **reciprocal** or **bilateral**; if the $z_{\kappa\lambda}$ or $y_{\kappa\lambda}$ are invariant under the change of the origin of time, it is said to be **time-invariant**; and if a linear time-invariant network satisfies the condition

$$\int_{-\infty}^t \sum_{B_\kappa \in S} i_\kappa(\tau) E_\kappa(\tau) d\tau \leq 0$$

for every t and for every choice of functions of time for i_κ or E_κ associated with the source branches B_κ in S , provided that the current-voltage relations for non-source branches are satisfied, then the network is said to be **passive**. Under certain nonsingularity conditions, for the currents in and the voltages across the source branches (denoted by I_κ and e_κ) of a linear M -port network, we have the relations

$$e_\kappa = \sum_{B_\lambda \in S} Z_{\kappa\lambda} I_\lambda, \quad I_\kappa = \sum_{B_\lambda \in S} Y_{\kappa\lambda} e_\lambda \quad (B_\kappa \in S),$$

where the matrices $Z_{\kappa\lambda}$ and $Y_{\kappa\lambda}$ of linear integrodifferential operators are called the **port-impedance matrix** and the **port-admittance matrix** of the network, respectively. **Analysis** determines $Z_{\kappa\lambda}$, $Y_{\kappa\lambda}$ from a given linear graph and given $z_{\kappa\lambda}$, $y_{\kappa\lambda}$, while **synthesis** finds a network (i.e., $z_{\kappa\lambda}$ or $y_{\kappa\lambda}$, as well as a linear graph) when part of $Z_{\kappa\lambda}$, $Y_{\kappa\lambda}$, or some relations to be satisfied by them are given. (In synthesis, the $z_{\kappa\lambda}$ or $y_{\kappa\lambda}$ are usually confined to some special class.) In analysis as well as synthesis we usu-

ally deal with the \dagger Laplace transforms $\bar{z}_{\kappa\lambda}(s)$, $\bar{y}_{\kappa\lambda}(s)$, $\bar{Z}_{\kappa\lambda}(s)$, $\bar{Y}_{\kappa\lambda}(s)$ instead of $z_{\kappa\lambda}$, $y_{\kappa\lambda}$, $Z_{\kappa\lambda}$, $Y_{\kappa\lambda}$ themselves, where the characteristics $Z_{\kappa\lambda}(s)$, $\bar{Y}_{\kappa\lambda}(s)$ of a network are determined by the topological properties of its linear graph and the properties of $\bar{z}_{\kappa\lambda}(s)$, $\bar{y}_{\kappa\lambda}(s)$ as functions of the complex variable s . (If ω is the angular frequency, $s = i\omega$.)

The following fundamental facts are known in regard to analysis and synthesis:

(1) If $\bar{Z}_{\kappa\lambda}(s)$, $\bar{Y}_{\kappa\lambda}(s)$ are rational functions of s holomorphic on the open right half-plane, a necessary and sufficient condition for a network to be passive is that for arbitrary real numbers ξ_{κ} ,

$$\sum_{B_{\kappa}, B_{\lambda} \in S} \xi_{\kappa} \xi_{\lambda} \bar{Z}_{\kappa\lambda}(s) \quad \text{or} \quad \sum_{B_{\kappa}, B_{\lambda} \in S} \xi_{\kappa} \xi_{\lambda} \bar{Y}_{\kappa\lambda}(s)$$

is a **positive real function** of s , i.e., a function whose value is real when s is real and lies on the right half-plane when s lies on the right half-plane.

(2) Every passive one-port network can be synthesized by using a finite number of three kinds of branches, i.e., positive resistors ($E_{\kappa} = R_{\kappa} i_{\kappa}$, $R_{\kappa} > 0$), positive capacitors ($i_{\kappa} = C_{\kappa} dE_{\kappa}/dt$, $C_{\kappa} > 0$), and positive inductors ($E_{\kappa} = L_{\kappa} di_{\kappa}/dt$, $L_{\kappa} > 0$).

(3) Every passive M -port network ($M \geq 2$) can be synthesized by using, in addition to the three kinds of branches mentioned in (2), ideal transformers (an ideal transformer is a pair of branches (B_{κ} , B_{λ}) such that $i_{\kappa} = n i_{\lambda}$, $E_{\lambda} = n E_{\kappa}$, and $n = \text{real number}$) and ideal gyrators (an ideal gyrator is a pair of branches (B_{κ} , B_{λ}) such that $i_{\kappa} = E_{\lambda}$, $i_{\lambda} = -E_{\kappa}$); ideal gyrators are not needed, however, to synthesize a reciprocal network.

However, very little is known about the synthesis of passive M -port networks without using ideal transformers and gyrators. Topological methods are expected to be powerful for such synthesis problems.

The linear graph structure of a network loses its significance if ideal transformers are admitted.

References

- [1] K. Kondo (ed.), RAAG memoirs of the unifying study of basic problems in engineering and the physical sciences by means of geometry, Association for Science Documents Information, Tokyo, I, 1955; II, 1958; III, 1962.
- [2] W. Cauer, Synthesis of linear communication networks, McGraw-Hill, second edition, 1958. (Original in German, 1941.)
- [3] J. E. Storer, Passive network synthesis, McGraw-Hill, 1957.
- [4] L. Weinberg, Network analysis and synthesis, McGraw-Hill, 1962.

- [5] O. Brune, Synthesis of a finite two-terminal network whose driving-point impedance is a prescribed function of frequency, J. Math. Phys., 10 (1931), 191–236.
- [6] Y. Oono and K. Yasuura, Synthesis of finite passive $2n$ -terminal networks with prescribed scattering matrices, Mem. Fac. Eng. Kyushu Univ., 14 (1954), 125–177.
- [7] S. Seshu and M. B. Reed, Linear graphs and electrical networks, Addison-Wesley, 1961.
- [8] W. H. Kim and R. T. Chien, Topological analysis and synthesis of communication networks, Columbia Univ. Press, 1962.
- [9] C. Berge and A. Ghouila-Houri, Programming, games and transportation networks, Methuen, 1965. (Original in French, 1962.)
- [10] L. R. Ford, Jr., and D. R. Fulkerson, Flows in networks, Princeton Univ. Press, 1962.
- [11] R. W. Newcomb, Linear multiport synthesis, McGraw-Hill, 1966.

283 (XXI.37) Newton, Isaac

Sir Isaac Newton (December 25, 1642–March 20, 1727), the English mathematician and physicist, was born into a family of farmers in Woolsthorpe, Lincolnshire. In 1661, he entered Cambridge University, where he was greatly influenced by the professor who was teaching geometry, I. Barrow, and where he began research in Kepler's optics and Descartes's geometry.

In 1665, he discovered the \dagger binomial theorem, and in the same year, during a stay at his birthplace to escape the plague, he began work on his three great discoveries—the spectral decomposition of light, the universal law of gravity, and differential and integral calculus. He returned to Cambridge University in 1667, and in the following year invented the reflecting telescope and proposed his theory of light particles. During this period, he succeeded Barrow as professor and lectured on optics. At the same time, he probed deeper into the calculus. Guided by Barrow's insight that differentiation and integration were inverse operations and also by his own research on infinite series, Newton obtained the \dagger fundamental theorem of calculus. Leibniz obtained the same theorem a little later, and a struggle resulted between the two over priority. The two discoveries were independent, but because Leibniz's notation was superior, the later development of calculus owes more to him. The dynamic elucidation of the heliocentric theory was accomplished in Newton's main

work, *Principia mathematica philosophiae naturalis* (1686–1687), in which Kepler's law on the movement of the planets, Galileo's theory of movement, and Huygens's theory of oscillation were unified into the three laws of Newtonian dynamics. These natural laws, which deal with all dynamic phenomena in the universe, are the most superlative realization of Descartes's concept of exploring the mathematical structure of nature; they had an essential influence on the later development of the natural sciences. The style of writing is similar to that in Euclid's *Stoicheia*. In the *Principia*, Newton also sets forth his philosophical position.

In 1695, Newton moved to London and became engrossed in theology. He was appointed Master of the Mint and was president of the Royal Society from 1703 until his death. While he is sometimes said to have divorced himself from science, many of his notes on geometry date from this time.

References

- [1] D. T. Whiteside (ed.), Sir Isaac Newton, Mathematical works, Johnson Reprint, I, 1964; II, 1969.
- [2] D. T. Whiteside (ed.), Sir Isaac Newton, Mathematical papers, Cambridge Univ. Press, I, 1967; II, 1968; III, 1969; IV, 1971; V, 1972; VI, 1974; VII, 1976; VIII, 1981.
- [3] I. Newton, *Philosophiae naturalis principia mathematica*, London, 1687; English translation, *Mathematical principles of natural philosophy*, trans. by A. Motte in 1729, Univ. of California Press, 1934.
- [4] D. Brewster, *Memoirs of the life, writings, and discoveries of Sir Isaac Newton*, Constable, 1855.

284 (III.12) Noetherian Rings

A. General Remarks

In this article, we mean by ring a commutative ring with unity element. Thus a **Noetherian ring** is a commutative ring with unity element that satisfies the \dagger maximum condition for its ideals; if it is also an \dagger integral domain, then it is called a **Noetherian integral domain** or **Noetherian domain** (for right and left Noetherian rings \rightarrow 368 Rings F).

A ring is Noetherian if and only if every prime ideal of the ring has a \dagger finite basis (**Cohen's theorem**). A ring is Noetherian if it is

generated by a finite number of elements over a Noetherian ring (**Hilbert's basis theorem**). The following three conditions for a ring R are equivalent: (i) R is an **Artinian ring**, that is, it satisfies the \dagger minimum condition for its ideals (for right and left Artinian rings \rightarrow 368 Rings F). (ii) R is a Noetherian ring and every prime ideal of R is a maximal ideal. (iii) There exist a finite number of Noetherian rings $R_i (i = 1, \dots, n)$ whose \dagger maximal ideals are \dagger nilpotent such that R is the direct sum of the $R_i (i = 1, \dots, n)$. We say that the **restricted minimum condition** holds in a ring R if R/\mathfrak{a} is an Artinian ring for every nonzero ideal \mathfrak{a} of R ; the latter condition is satisfied if and only if R is either an Artinian ring or a Noetherian domain of \dagger Krull dimension 1. Every ideal of a Noetherian ring R can be expressed as the intersection of a finite number of \dagger primary ideals. Given a ring R and an R -module M , a submodule P of M is said to be a **primary submodule** of M if every element a of R that is a **zero divisor with respect to M/P** (i.e., there exists an $m \in M/P$ such that $m \neq 0$ and $am = 0$) is **nilpotent with respect to M/P** (i.e., there exists a natural number n such that $a^n(M/P) = 0$).

The property of ideals in Noetherian rings stated above can be generalized to the case of Noetherian modules: If an R -module M is a \dagger Noetherian module, then every submodule of M can be expressed as the intersection of a finite number of primary submodules.

Let R be a Noetherian ring, \mathfrak{a} an ideal of R , M a finite R -module, and N and N' submodules of M . Then we have (i) the **Artin-Rees lemma**: There exists a natural number r such that for all $n > r$, $\mathfrak{a}^n N \cap N' = \mathfrak{a}^{n-r} \cdot (\mathfrak{a}^r N \cap N')$. (ii) **Krull's intersection theorem**: $\bigcap_{n=1}^{\infty} \mathfrak{a}^n M = \{m \in M \mid \exists a \in \mathfrak{a} \text{ such that } (1-a)m = 0\}$ (hence, in particular, if \mathfrak{m} is the \dagger Jacobson radical of R , then $\bigcap_{n=1}^{\infty} \mathfrak{m}^n M = \{0\}$). (iii) **Krull's altitude theorem**: If \mathfrak{a} is generated by s elements and \mathfrak{p} is a \dagger minimal prime divisor of \mathfrak{a} , then the height of $\mathfrak{p} \leq s$.

B. Topology Defined by an Ideal

Let R be a ring, \mathfrak{a} an ideal of R , and M an R -module. Then the **\mathfrak{a} -adic topology** of M is defined to be the topology on M such that $\{\mathfrak{a}^n M \mid n = 1, 2, \dots\}$ is a \dagger base for the neighborhood system of zero. In particular, let R be Noetherian, M a finite R -module, and N a submodule of M . Then by the Artin-Rees lemma, the \mathfrak{a} -adic topology of N coincides with the topology on N as a subspace of M with the \mathfrak{a} -adic topology. Returning to the general case, M is a $\dagger T_0$ -space (under the \mathfrak{a} -adic topology) if and only if $\bigcap_{n=1}^{\infty} \mathfrak{a}^n M = \{0\}$.

or, in other words, if and only if M is a \dagger metric space, where the \dagger distance $d(a, b)$ between points a, b in M is defined to be $\inf\{2^{-n} \mid a - b \in \alpha^n M\}$. Moreover, if N is a submodule of M , then M/N is a T_0 -space if and only if N is a closed subset of M , that is, if and only if $\bigcap_{n=1}^{\infty} (N + \alpha^n M) = N$. A sequence $\{a_n\} = (a_1, a_2, \dots, a_n, \dots)$ is called a **Cauchy sequence** (under the α -adic topology) if $\forall n \exists N \forall r \forall s$ (with each of them a natural number) $a_{N+r} - a_{N+s} \in \alpha^n M$; for this it is sufficient that $\forall n \exists N \forall r \ a_{N+r+1} - a_{N+r} \in \alpha^n M$. If this sequence converges to zero (i.e., $\forall n \exists N \forall r \ a_{N+r} \in \alpha^n M$), it is called a **null sequence**. The set \mathfrak{M} of all Cauchy sequences in M becomes an R -module if we define their sum and multiplication by an element of R by $\{a_n\} + \{b_n\} = \{a_n + b_n\}$, $c\{a_n\} = \{ca_n\}$. Then the set \mathfrak{N} of all null sequences is a submodule of \mathfrak{M} . An element m of M is identified with the sequence (m, m, \dots, m, \dots) , and we regard M as a submodule of \mathfrak{M} . Then the R -module $\bar{M} = \mathfrak{M}/\mathfrak{N}$ is the \dagger completion of $M/(\bigcap_{n=1}^{\infty} \alpha^n M)$ (as a metric space under the α -adic topology). \bar{M} is called the **α -adic completion** of M . If α has a finite basis, then the topology of \bar{M} (as the completion) coincides with its α -adic topology. If $M = R$, we define a multiplication in \mathfrak{M} by $\{a_n\}\{b_n\} = \{a_n b_n\}$. In this case, \mathfrak{M} is a ring in which \mathfrak{N} is an ideal, and hence the completion $\bar{R} = \bar{M}$ is a ring. If $\alpha = \sum_{i=1}^r a_i R$, then considering the \dagger ring of formal power series $\tilde{R} = R[[x_1, \dots, x_r]]$ and its ideal $\tilde{n} = \bigcap_{n=1}^{\infty} (\sum_{i=1}^r (a_i - x_i) \tilde{R} + \alpha^n \tilde{R})$, we have $\bar{R} \cong \tilde{R}/\tilde{n}$.

C. Zariski Rings

For a Noetherian ring R and an ideal α of R , every element b of R such that $1 - b \in \alpha$ has an inverse in R if and only if every ideal of R is a closed subset of R under the α -adic topology. When this condition is satisfied, the ring R with the α -adic topology is called a **Zariski ring**; we often express this by saying that (R, α) is a Zariski ring. A Zariski ring is called **complete** if it is a complete topological space. The completion \bar{R} of the Zariski ring (R, α) has the $\alpha\bar{R}$ -adic topology, and $(\bar{R}, \alpha\bar{R})$ is a Zariski ring. Furthermore, (i) \bar{R} is a \dagger faithfully flat R -module; and (ii) when N is a submodule of a finite R -module M , then N is a closed subspace of M (under their α -adic topologies), and their completions are identified with $N \otimes_R \bar{R}$, $M \otimes_R \bar{R}$.

D. Local Rings

Suppose that R is a Noetherian ring having only a finite number of maximal ideals and J is

the Jacobson radical of R . Then the Zariski ring (R, J) is called a **semilocal ring**. Furthermore, if R has only one maximal ideal, then (R, J) is called a **local ring**. A ring that has only a finite number of maximal ideals is called a **quasisemilocal ring**; if it has only one maximal ideal, it is called a **quasilocal ring**. (In some literature, the terms *local ring* and *semilocal ring* are used under weaker conditions; in the weakest case, quasisemilocal rings and quasilocal rings are simply called semilocal rings and local rings, respectively, and local rings and semilocal rings in our sense are called **Noetherian local rings** and **Noetherian semilocal rings**, respectively.)

Assume that R is a semilocal ring with maximal ideals $\mathfrak{m}_1, \dots, \mathfrak{m}_n$ and the Jacobson radical $J = \mathfrak{m}_1 \cap \dots \cap \mathfrak{m}_n$. For every finite R -module M , we introduce the J -adic topology as its natural topology. The completion \bar{R} of R is a semilocal ring with maximal ideals $\mathfrak{m}_1 \bar{R}, \dots, \mathfrak{m}_n \bar{R}$ and is naturally isomorphic to the direct sum of the completions of the local rings $R_{\mathfrak{m}_i}$ ($i = 1, \dots, n$). Since R is a Zariski ring, (1) \bar{R} is faithfully flat; (2) submodules of M are closed subsets of M ; and (3) the completion of M is identified with $M \otimes_R \bar{R}$. If (R, \mathfrak{m}) is a **complete local ring** (i.e., a local ring and a complete Zariski ring at the same time), then R contains a subring I with the following properties: (i) I is a complete local ring, and $I/(\mathfrak{m} \cap I) = R/\mathfrak{m}$; and (ii) for the \dagger characteristic p of R/\mathfrak{m} (p is either zero or a prime number), $\mathfrak{m} \cap I = pI$. Therefore, if \mathfrak{m} is generated by n elements, then R is a homomorphic image of the ring of formal power series in n variables over I . This theorem is called the **structure theorem of complete local rings**, and I is called a **coefficient ring** of R . If R contains a field, then I is a field, called a **coefficient field** of R . A complete local ring is a \dagger Hensel ring.

When (R, \mathfrak{m}) is a local ring and $\sum_{i=1}^r x_i R$ is \mathfrak{m} -primary, then we have an inequality $r \geq (\dagger \text{Krull dimension of } R)$; if the equality holds, then we say that x_1, \dots, x_r form a **system of parameters** of R . Furthermore, if $\mathfrak{m} = \sum_{i=1}^r x_i R$, then we say that x_1, \dots, x_r form a **regular system of parameters** of R . A local ring that has a regular system of parameters is called a **regular local ring** (cf. \dagger Jacobian criterion). A regular local ring is a \dagger unique factorization domain. Let d be the Krull dimension of a local ring (R, \mathfrak{m}) . Then R is a regular local ring if and only if one of the following holds: (1) every R -module has finite \dagger homological dimension, (2) every R -module has homological dimension of at most d , or (3) the homological dimension of R/\mathfrak{m} (as an R -module) is finite (and actually coincides with d). A Noetherian ring R' is called a **regular ring** if R'_p is

a regular local ring for every prime ideal \mathfrak{p}' . A regular local ring is a regular ring.

Consider a local ring (R, \mathfrak{m}) and an \mathfrak{m} -primary ideal \mathfrak{q} . Then the length $l(n)$ of R/\mathfrak{q}^n (as an R -module) is a function of n . For a sufficiently large n , the length $l(n)$ can be expressed as a polynomial $f(n)$ in n with rational coefficients. The degree of $f(n)$ coincides with the Krull dimension d of R . The **multiplicity** of \mathfrak{q} is (coefficient of n^d in $f(n)$) $\times (d!)$. If x_1, \dots, x_d form a system of parameters, then (multiplicity of $\sum_{i=1}^d x_i R$) \leq (length of $R/\sum_{i=1}^d x_i R$); if the equality holds, then we call x_1, \dots, x_d a **distinct system of parameters**. A local ring that has a distinct system of parameters is called a **Macaulay local ring**. A local ring is a Macaulay local ring if and only if one of the following holds: (1) every system of parameters is a distinct system of parameters, or (2) if an ideal \mathfrak{a} of height s is generated by s elements, then every \mathfrak{p} -prime divisor of \mathfrak{a} is of height s . A regular local ring is a Macaulay local ring.

The notion of multiplicity can be also defined in general Noetherian rings [4]. Let R be a Noetherian ring. If $R_{\mathfrak{m}}$ is a Macaulay local ring for every maximal ideal \mathfrak{m} , then R is called a **locally Macaulay ring**. Furthermore, if height $\mathfrak{m} = \text{Krull dim } R$ for every maximal ideal \mathfrak{m} , then R is called a **Macaulay ring**. If R is a locally Macaulay ring, then the polynomial ring in a finite number of variables over R is also a locally Macaulay ring. In general, an ideal \mathfrak{a} is called an **unmixed ideal** (or **pure ideal**) if the height of every prime divisor of \mathfrak{a} coincides with height \mathfrak{a} ; otherwise, \mathfrak{a} is called a **mixed ideal**. Thus if R is a locally Macaulay ring, an ideal \mathfrak{a} of the polynomial ring $R[x_1, \dots, x_n]$ over R is generated by r elements, and height $\mathfrak{a} = r$, then \mathfrak{a} is unmixed (**unmixedness theorem**).

If the completion of a local ring R is a \mathfrak{p} -normal ring, then we say that R is **analytically normal**. If the completion of a semilocal ring R has no nilpotent element except zero, then we say that R is **analytically unramified**. A semilocal integral domain R which is a ring of quotients of a finitely generated ring over a field is analytically unramified. If R is a normal local ring, then R is analytically normal (O. Zariski).

Let (R, \mathfrak{m}) be a local ring, and let \mathfrak{q} be an \mathfrak{m} -primary ideal. Set $F_i = \mathfrak{q}^i / \mathfrak{q}^{i+1}$ ($i = 0, 1, 2, \dots$, $\mathfrak{q}^0 = R$). Let $a = a' \pmod{\mathfrak{q}^{i+1}} \in F_i$ and $b = b' \pmod{\mathfrak{q}^{j+1}} \in F_j$. We put $ab = a'b' \pmod{\mathfrak{q}^{i+j+1}} \in F_{i+j}$. Then the direct sum of modules $F = \sum_{i=0}^{\infty} F_i$ becomes a graded ring generated by F_1 over F_0 , in which F_i is the module of homogeneous elements of degree i . F , called the **form ring** (or **associated graded ring**) of R with respect to \mathfrak{q} , plays an important role in the

theory of local rings, particularly in the theory of multiplicity.

E. Chains of Prime Ideals

Let R be a Noetherian ring with prime ideals $\mathfrak{p}, \mathfrak{q}$ such that $\mathfrak{p} \subset \mathfrak{q}$. Consider the length n of a chain of prime ideals $\mathfrak{p} = \mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \dots \subsetneq \mathfrak{p}_n = \mathfrak{q}$ which cannot be refined any more. It is not true in general that n is uniquely determined by \mathfrak{p} and \mathfrak{q} (M. Nagata). However, n is uniquely determined for a rather large class of Noetherian rings, for instance the rings that are homomorphic images of locally Macaulay rings and, in particular, finitely generated rings over a \mathfrak{p} -Dedekind domain.

F. Integral Closures

Let R be a Noetherian integral domain with the field of quotients k , let K be a finite algebraic extension of k , and let \tilde{R} be the \mathfrak{p} -integral closure of R in K . Then (i) If R is of Krull dimension 1, then for an arbitrary ring R' such that $R \subset R' \subset K$ and for every nonzero ideal \mathfrak{a}' of R' , the quotient R'/\mathfrak{a}' is a finite $R/(\mathfrak{a}' \cap R)$ -module. In particular, R' is a Noetherian domain satisfying the restricted minimum condition. (ii) If R is of Krull dimension 2, then \tilde{R} is Noetherian. (iii) In general, \tilde{R} is a \mathfrak{p} -Krull ring, and for an arbitrary prime ideal \mathfrak{p} of R there are only a finite number of prime ideals $\tilde{\mathfrak{p}}$ of \tilde{R} such that $\mathfrak{p} = \tilde{\mathfrak{p}} \cap R$. For any of such $\tilde{\mathfrak{p}}$, the field of quotients of $\tilde{R}/\tilde{\mathfrak{p}}$ is a finite algebraic extension of that of R/\mathfrak{p} . Result (i) is called the **Krull-Akizuki theorem**. We say that R satisfies the **finiteness condition for integral extensions** if \tilde{R} is a finite R -module for any choice of K . A Noetherian ring R is called a **pseudogeometric ring**, or a **universally Japanese ring**, if R/\mathfrak{p} satisfies the finiteness condition for integral extensions for every prime ideal \mathfrak{p} . A ring is pseudogeometric if it is generated by a finite number of elements over a pseudogeometric ring.

G. History

J. W. R. Dedekind first introduced the concept of ideals in the theory of integers. The main objects studied in ring theory were subrings of number fields or function fields until M. Sono (*Mem. Coll. Sci. Univ. Kyoto*, 2 (1917), 3 (1918–1919)) originated an abstract study of Dedekind domains, which was followed by E. Noether (*Math. Ann.*, 83 (1921), 96 (1926)), who originated the theory of Noetherian rings.

W. Krull made further important contributions to the development of the theory of Noetherian rings and general commutative rings [1]. Many other authors, including E. Artin, Y. Akizuki, and S. Mori, also contributed to the theory. The theory of local rings was originated by Krull (*J. Reine Angew. Math.*, 179 (1938)) and developed by C. Chevalley (*Ann. Math.*, 44 (1943)), I. S. Cohen (*Trans. Amer. Math. Soc.*, 59 (1946)), and Zariski (*Ann. Inst. Fourier*, 2 (1950)), and later by many authors, including P. Samuel, Nagata, M. Auslander, D. A. Buchsbaum, and J.-P. Serre [4]. The theory of Noetherian rings is applied to algebraic geometry.

References

- [1] W. Krull, *Idealtheorie*, Erg. Math., Springer, second edition, 1968.
- [2] B. L. van der Waerden, *Algebra I, II*, Springer, 1966–1967.
- [3] O. Zariski and P. Samuel, *Commutative algebra I, II*, Van Nostrand, 1958–1960.
- [4] M. Nagata, *Local rings*, Wiley, 1962 (Krieger, 1975).
- [5] N. Bourbaki, *Eléments de mathématique*, Algèbre, ch. 1, 8, *Actualités Sci. Ind.*, 1144b, 1261a, Hermann, 1964, 1958.
- [6] N. Bourbaki, *Eléments de mathématique*, Algèbre commutative, ch. 1–7, *Actualités Sci. Ind.*, Hermann; ch. 1, 2, 1290a, 1961; ch. 3, 4, 1293a, 1967; ch. 5, 6, 1308, 1964; ch. 7, 1314, 1965.
- [7] J.-P. Serre, *Algèbre locale, multiplicités*, Lecture notes in math. 11, Springer, 1965.
- [8] D. G. Northcott, *Lessons on rings, modules and multiplicities*, Cambridge Univ. Press, 1968.
- [9] H. Matsumura, *Commutative algebra*, Benjamin, 1970; second edition, 1980.

285 (VI.16) Non-Euclidean Geometry

A. History

The validity of the fifth postulate of Euclid's *Elements*, the \dagger axiom of parallels, has been a subject of argument ever since it was formulated (\rightarrow 139 Euclidean Geometry). At the beginning of the 18th century, G. Saccheri tried to prove the postulate by assuming the validity of other axioms. Under the hypothesis that the axiom does not hold, he deduced various extraordinary results. Although he

was mistaken in thinking that he had obtained a contradiction, his work is regarded as a forerunner to the study of **non-Euclidean geometry**.

At the beginning of the 19th century, N. I. Lobachevskii and J. Bolyai opened up the impasse by establishing a geometry based on postulates that contradict the fifth postulate. This geometry is called **hyperbolic geometry** or **Lobachevskii's non-Euclidean geometry**. Actually, a similar idea had been conceived by C. F. Gauss, but he refrained from publishing it because of likely misunderstanding by a public still strongly influenced by I. Kant's philosophy. On the other hand, B. Riemann constructed so-called **elliptic geometry** (or **Riemann's non-Euclidean geometry**), which is different from both Euclidean and hyperbolic geometry. Euclidean geometry (including the theory of similarity) is sometimes called **parabolic geometry**. In general, a space satisfying axioms that contradict Euclid's postulates is called a **non-Euclidean space**.

Around the turn of the 20th century, A. Cayley, F. Klein, and H. Poincaré constructed models of non-Euclidean spaces that are subsets of Euclidean spaces, and E. Beltrami constructed a differential geometric model. By means of these models, it was established that non-Euclidean geometries are consistent as long as there is no inconsistency in the underlying Euclidean geometries. On the other hand, D. Hilbert established a complete system of axioms for Euclidean geometry and showed, by constructing non-Euclidean models, that the axiom of parallels is independent of the other axioms (\rightarrow 155 Foundations of Geometry). The logical foundation of non-Euclidean geometries was thus clarified. Moreover, A. Einstein showed in his \dagger theory of relativity that actual space-time does not satisfy Euclidean axioms. Together with Euclidean spaces, non-Euclidean spaces are often used as fundamental models both in the problem of \dagger space forms and in the theory of \dagger symmetric spaces.

B. Axiomatic Considerations

Hilbert's system of axioms of plane Euclidean geometry consists of axioms of incidence, order, congruence, parallels, and continuity (\rightarrow 155 Foundations of Geometry). Specifically, the axiom of parallels is stated as follows: Suppose we are given a straight line and a point in a plane. If the straight line does not contain the point, then there exists only one line through the point that does not intersect the line.

The system of axioms of hyperbolic geometry is obtained by replacing the axiom of parallels by the following: Suppose we are given a straight line and a point in a plane. If the line does not contain the point, then there exist at least two lines that pass through the point without intersecting the line. (The other four groups of axioms are unaltered.) In this case, if l is a given line that does not contain a given point C , then there exist exactly two lines **parallel** to l that pass through C . We denote them by XY and $X'Y'$, and they are characterized as follows (Fig. 1): Any line that passes through C and lies in $\angle X'CY$ necessarily intersects l ; by contrast, neither the two lines XY , $X'Y'$ nor any line in $\angle XCY$ intersects l . Euclidean geometry can be considered as a "limit" of this geometry where the lines XY and $X'Y'$ coincide.

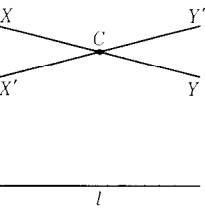


Fig. 1

In elliptic geometry, the axiom of parallels is replaced by the following: Suppose we are given a straight line and a point in a plane. If the line does not contain the point, then any line passing through the point intersects the line. In this case, the lines are closed curves, and the axioms of order must be modified. Specifically, in Euclidean geometry, the axioms of order are based on the notion of a point A lying between points B and C , where A, B, C are distinct points on a line. In elliptic geometry, however, to define order we utilize the notion of a pair A, C of points separating another pair B, D (and vice versa), where A, B, C, D are distinct points on a line. The axioms of order are modified accordingly.

The sum of inner angles of a triangle is smaller or greater than two right angles according as we use the axioms of hyperbolic or elliptic geometry.

C. The Projective-Geometric Point of View

We take \dagger projective coordinates in an n -dimensional, real \dagger projective space \mathbf{P}^n and consider a quadric hypersurface defined by $Q: ax_0^2 + x_1^2 + \dots + x_n^2 = 0, a \neq 0$ (\rightarrow 90 Coordinates; 343 Projective Geometry). We denote by G the group consisting of the totality of projective transformations of \mathbf{P}^n that leave Q

invariant. We call Q the **absolute** and call G the **group of congruent transformations**. When $a < 0$, then this Q is a real quadric hypersurface. In this case, there exists a domain H^n (the totality of points inside Q) whose boundary coincides with Q , and the group G acts \dagger transitively on H^n . The pair $\{G, H^n\}$ provides a model of hyperbolic geometry, and the n -dimensional **hyperbolic space** H^n is homeomorphic to an n -dimensional open cell. Points of H^n , points on Q , and points outside Q are called **ordinary points**, **points at infinity**, and **ultrafinite points** (or **ideal points**), respectively. Two lines on H^n are said to be parallel if they intersect on the absolute Q . Next, when $a > 0$, the absolute Q is an imaginary quadric hypersurface, and the group G acts transitively on \mathbf{P}^n . The pair $\{G, \mathbf{P}^n\}$ provides a model of elliptic geometry. The n -dimensional **elliptic space** \mathbf{P}^n is homeomorphic to n -dimensional real projective space; hence it is compact. In elliptic geometry, any two distinct lines in a plane necessarily intersect at a point. The above models $\{G, \mathbf{P}^n\}$ of non-Euclidean geometries are called **Klein's models**.

Let $\{G, H^n\}$ be a Klein's model, A, B distinct points in \mathbf{P}^n , l the line containing A, B , and I, J be two points where the line l meets Q (Fig. 2). If we denote by (A, B, I, J) the \dagger anharmonic ratio of these four points, then the **non-Euclidean distance** ρ between the points A and B is given by $\rho = a \log(A, B, I, J)$, a a constant. Next, let l, g be lines in H^n intersecting at a point D . In the plane determined by l and g , we draw two imaginary tangents u, v to the absolute through D . Denoting by (l, g, u, v) the anharmonic ratio of these four lines, the **non-Euclidean angle** θ between the lines l and g is given by $\theta = (1/2i) \log(l, g, u, v), i = \sqrt{-1}$. Generally, let O be a point in the projective space \mathbf{P}^n , and denote by p the \dagger polar of O with respect to the absolute. By counting the polar p doubly, it can be regarded as a quadric hypersurface, which will be denoted by S_0 . A quadric hypersurface S of H^n is called a **non-Euclidean hypersphere** if it belongs to the \dagger pencil of quadric hypersurfaces determined by Q and S_0 . S is called a **proper hypersphere**, a **limiting hypersphere**, or an **equidistant hypersurface** according as the center O is an ordi-

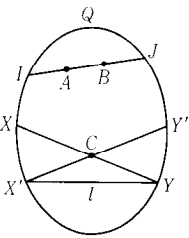


Fig. 2

nary point, a point at infinity, or an ultrainfinite point, respectively. In the case of elliptic geometry $\{G, \mathbf{P}^n\}$, the distance ρ between two points A and B is given by $\rho = (a/i) \cdot \log(A, B, I, J)$ as I, J are imaginary points. Moreover, we may get parabolic geometry as the "limit" ($a \rightarrow \infty$) of the geometry given by Klein's models.

D. The Conformal-Geometric Point of View

Let S^n be an n -dimensional † conformal space. If we take suitable $(n+2)$ -hyperspherical coordinates in S^n , the space S^n can be realized as a quadric hypersurface $x_1^2 + x_2^2 + \dots + x_n^2 - 2x_0x_\infty = 0$ in $(n+1)$ -dimensional real projective space \mathbf{P}^{n+1} . A point in \mathbf{P}^{n+1} represents a hypersphere in S^n (\rightarrow 76 Conformal Geometry; 90 Coordinates). We denote by \tilde{G} the group consisting of the totality of † Möbius transformations leaving invariant a (real, point, or imaginary) hypersphere Q . When Q is a real hypersphere, the space S^n is divided into Q and the two open cells H^n, H_∞^n . If we denote by G the totality of transformations of the group \tilde{G} that do not interchange H^n and H_∞^n , then G is a subgroup of index 2 of \tilde{G} . In this case, each of the pairs $\{G, H^n\}$ and $\{G, H_\infty^n\}$ provides a model of hyperbolic geometry. When Q is a point hypersphere, the space E^n obtained from S^n by omitting the point Q is homeomorphic to an open cell, and the pair $\{\tilde{G}, E^n\}$ provides a model of parabolic geometry. On the other hand, when Q is an imaginary hypersphere, \tilde{G} is isomorphic to the † orthogonal group $O(n+1)$. We call the pair $\{\tilde{G}, S^n\}$ a **spherical geometry** and S^n an n -dimensional **spherical space**. In this case two points x and x' are called equivalent if x' is the image of x by symmetry with respect to Q . The space P^n obtained from S^n by identifying equivalent points is homeomorphic to the n -dimensional real projective space. If we denote by G the group obtained from \tilde{G} by making its actions effective on P^n , then G is the factor group of \tilde{G} by a cyclic group $\mathbf{Z}_2 = \mathbf{Z}/2\mathbf{Z}$. The pair $\{G, P^n\}$ provides a model of elliptic geometry. These models are called **Poincaré's models**. They were introduced as a result of research on automorphic functions in the case $n=2$.

In Poincaré's model, every straight line is represented either by a circle orthogonal to Q or by a circle passing through Q according as Q is a (real or imaginary) hypersphere or a point hypersphere. In spherical geometry, however, straight lines are usually called great circles, and two distinct great circles lying on a 2-dimensional sphere necessarily intersect at two points that are symmetric with respect to Q . Also, in Poincaré's model the distance

between two points A and B is defined as before, by making use of the anharmonic ratio of four points A, B, I, J on a circle (Fig. 3) (\rightarrow 74 Complex Numbers G).

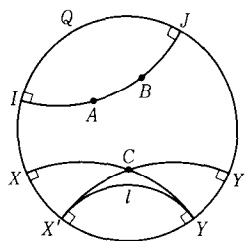


Fig. 3

E. The Differential-Geometric Point of View

An n -dimensional space M of † constant curvature is by definition a † Riemannian manifold whose line element ds is given by

$$ds^2 = \frac{dx_1^2 + dx_2^2 + \dots + dx_n^2}{\left(1 + \frac{K}{4}(x_1^2 + x_2^2 + \dots + x_n^2)\right)}$$

with respect to appropriate local coordinates, where K is a constant called the † sectional curvature (\rightarrow 364 Riemannian Manifolds). According as K is positive, zero, or negative, M can be considered locally as an elliptic space, Euclidean space, or hyperbolic space, respectively. In this case, lines are † geodesics of M , and the non-Euclidean distance and non-Euclidean angle are those defined in the Riemannian manifold. When $n=2$, a † simply connected and † complete space of positive constant curvature is † embedded in 3-dimensional Euclidean space as a sphere, and a space of negative constant curvature is † locally isometric to a **pseudosphere** (Fig. 4), which is a surface of revolution obtained by rotating a † tractrix around its asymptote. † Complete n -dimensional spaces of constant curvature ($n \geq 2$) are called **space forms**. A simply connected space form is necessarily one of spherical space, Euclidean space, or hyperbolic space. Each of these is a † universal covering manifold of a general connected space form with a curvature of the same sign, and the group of † covering transformations is isomor-

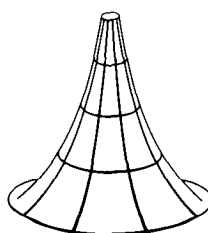


Fig. 4

phic to a †discontinuous subgroup of the group of congruent transformations.

Each of the spaces, Euclidean, non-Euclidean, and spherical, is a †homogeneous space on which the corresponding group of congruent transformations acts transitively. Actually, each of these spaces has the structure of a †symmetric Riemannian homogeneous space of rank 1.

References

- [1] D. M. Y. Sommerville, *The elements of non-Euclidean geometry*, Bell, 1914.
- [2] F. Klein, *Vorlesungen über nicht-Euklidische Geometrie*, Springer, 1928 (Chelsea, 1960).
- [3] E. Cartan, *Leçons sur la géométrie des espaces de Riemann*, Gauthier-Villars, second edition, 1963.
- [4] H. S. M. Coxeter, *Non-Euclidean geometry*, Univ. of Toronto Press, fifth edition, 1965.
- [5] F. Engel and P. Stäckel, *Urkunden zur Geschichte der nicht-Euklidischen Geometrie I, II*, Teubner, 1898, 1913.

286 (XII.21) Nonlinear Functional Analysis

A. General Remarks

At present, the theory of nonlinear problems is still not unified, and many individual results obtained for specific classes of problems are stated in the languages of the corresponding fields of study. However, there are some fundamental facts and methods of a general nature concerning nonlinear problems, which may be referred to as the subject matter of **nonlinear functional analysis**.

B. Iterative Methods

Let X be a †Banach space. Consider a nonlinear mapping $G: X \rightarrow X$ and the equation

$$Gx = 0 \quad (x \in X). \quad (1)$$

Set $F = I - G$. Then (1) can be written as

$$x = Fx. \quad (2)$$

F satisfies the **Lipschitz condition** if there exists a constant α such that

$$\|Fx - Fy\| \leq \alpha \|x - y\| \quad (3)$$

for all x, y in X . In particular, the mapping F is said to be **nonexpansive** if $0 < \alpha \leq 1$. If $0 < \alpha < 1$, then F is called a **contraction**. (Sometimes a nonexpansive mapping is also called a contraction.) A contraction F satisfies the **contraction principle**: F has a unique fixed point x_0 , and the **iteration**

$$x_{n+1} = Fx_n \quad (n = 1, 2, \dots)$$

with an arbitrary initial element x_1 always converges to x_0 [1–3]. Similar results hold for a contraction that is defined only on a †ball and leaves the ball invariant. This leads to †Newton's iterative process and the †implicit function theorem.

C. Methods of Monotonicity

By definition, a nonlinear mapping G from a †Hilbert space H to H is a **monotone** or **accretive operator** if

$$\operatorname{Re}(Gx - Gy, x - y) \geq 0 \quad (x, y \in H).$$

G. Minty [4] proved that if $G: H \rightarrow H$ is monotone and continuous, then $\lambda I + G$ is a mapping onto H for any $\lambda > 0$, and its inverse $(\lambda I + G)^{-1}$ is nonexpansive. He has also shown that in the hypothesis of the theorem, we can replace the continuity requirement for G by maximality of G within the class of accretive operators that are possibly multivalued. Various developments of Minty's ideas, including generalization of his results to Banach spaces and applications to partial differential equations, have been obtained by F. Browder (*Amer. Math. Soc. Proc. Symposia in Appl. Math.*, 17 (1965)), J. Leray and J. L. Lions (*Bull. Soc. Math. France*, 93 (1965)), J. L. Lions [5], W. Strauss, H. Brézis (*Amer. Math. Soc. Proc. Symposia in Pure Math.*, 18 (1970)), and others.

A mapping A is said to be **dissipative** if $-A$ is accretive. Dissipative mappings play a central role in the theory of nonexpansive semi-groups (\rightarrow Section X).

D. Topological Methods

In the geometric study of ordinary differential equations [6] some familiar theorems of topology and †differential topology have been strong tools, e.g., †Brouwer's fixed-point theorem is utilized to establish the existence of periodic solutions. However, in order to deal with nonlinear partial differential equations we have to generalize these theorems to infinite-dimensional cases. For example, a fixed-point theorem in an infinite-dimensional space was first obtained by G. D. Birkhoff and O. D.

Kellogg (*Trans. Amer. Math. Soc.*, 23 (1922), 95–115). The theory of the degree of mappings was generalized to the case of Banach spaces by J. Leray and J. Schauder [33] for the class of mappings of the form $I - F$, where F is a **compact** continuous mapping, i.e., the image by F of any bounded set is relatively compact. Let D be an open bounded set in a Banach space X and ∂D be its boundary. Let $F: \bar{D} \rightarrow X$ be a continuous compact mapping and Φ denote $I - F$. If a point p in X does not belong to $\Phi(\partial D)$, then we can define the **Leray-Schauder degree** $d(\Phi, p, D)$ of Φ relative to p [1–3]. The Leray-Schauder degree $d(\Phi, p, D)$ is an integer with the following properties: (i) $d(I, p, D) = 1$ if $p \in D$. If $p \notin \Phi(D)$, then $d(\Phi, p, D) = 0$. (ii) (Homotopy invariance) $d(\Phi, p, D)$ depends only on the compact homotopy class of $\Phi: \partial D \rightarrow X \setminus \{p\}$. More precisely, let $K: [0, 1] \times \partial D \rightarrow X$ be a continuous compact mapping such that $x + K(t, x) \neq p$ for any $t \in [0, 1]$ and any $x \in \partial D$. Let $\Phi_0(x) = x + K(0, x)$ and $\Phi_1(x) = x + K(1, x)$. Then $d(\Phi_0, p, D) = d(\Phi_1, p, D)$. (iii) If p and p' are in the same component of $X \setminus \Phi(\partial D)$, then $d(\Phi, p, D) = d(\Phi, p', D)$. (iv) (Continuity) $d(\Phi, p, D)$ is a continuous locally constant function of Φ (with respect to uniform convergence) and of $p \in X \setminus \Phi(\partial D)$. (v) (Domain decomposition) If D is the union of finite number of open disjoint sets D_j ($j = 1, 2, \dots, N$) with $\partial D_j \subset \partial D$ and $\Phi(x) \neq p$ on $\bigcup_{j=1}^N \partial D_j$, then $d(\Phi, p, D) = \sum_{j=1}^N d(\Phi, p, D_j)$. (vi) (Excision) If Δ is a closed subset of D on which $\Phi(x) \neq p$, then $d(\Phi, p, D) = d(\Phi, p, D \setminus \Delta)$. (vii) (Cartesian product formula) If $X = X_1 \oplus X_2$ with $D_i \subset X_i$, $\Phi = (\Phi_1, \Phi_2)$ with $\Phi_i: \bar{D}_i \rightarrow X_i$ ($i = 1, 2$), $D = D_1 \times D_2$, and $p = (p_1, p_2)$, then $d(\Phi, p, D) = d(\Phi_1, p_1, D_1) \times d(\Phi_2, p_2, D_2)$, provided that the right-hand side is well defined.

The degree of mapping of Φ is also defined for some proper Fredholm mapping Φ (\rightarrow Section E; K. D. Elworthy and A. J. Tromba [38]; [32]).

Brouwer's fixed-point theorem (\rightarrow 153 Fixed-Point Theorems) is generalized to **Schauder's fixed-point theorem**: A compact mapping F of a closed bounded convex set K in a Banach space X into itself has a fixed point. Using the Leray-Schauder degree theory, one has the **Leray-Schauder fixed-point theorem**: Let D be a bounded open set of a Banach space X containing the origin O . Let $F(x, t): \bar{D} \times [0, 1] \rightarrow X$ be a compact mapping such that $F(x, 0) \equiv 0$. Suppose that $F(x, t) \neq x$ for any $x \in \partial D$ and $t \in [0, 1]$. Then the compact mapping $F(x, 1)$ has a fixed point x in D [1–3].

Other homotopy invariants, such as * homotopy groups and * cohomotopy groups, are also used in nonlinear functional analysis [1, 3]. For instance we have the following

theorem of Shvarts (1964) [3]: Let X and Y be two Banach spaces. Let $D = \{x \in X \mid \|x\| < 1\}$ be the unit ball and $\partial D = \{x \in X \mid \|x\| = 1\}$ be the unit sphere of X . Suppose L is a fixed continuous linear * Fredholm operator from X to Y of index $p \geq 0$. Let PL be the set of compact perturbation of L mapping ∂D into $Y \setminus 0$, i.e., $PL = \{\Phi = L + K \mid K \text{ is a continuous compact mapping of } \partial D \text{ to } Y \text{ such that } \Phi(x) = Lx + K(x) \neq 0 \text{ for } x \in \partial D\}$. Two mappings Φ_0 and Φ_1 in PL are said to belong to the same **compact homotopy class** on ∂D if there exists a continuous compact mapping $h: [0, 1] \times \partial D \rightarrow Y$ such that $Lx + h(t, x) \neq 0$ for x in ∂D , $\Phi_0(x) = Lx + h(0, x)$, and $\Phi_1(x) = Lx + h(1, x)$.

Shvarts's theorem: Let L be a fixed continuous linear Fredholm mapping from X to Y of index $p \geq 0$. Then the compact homotopy classes on ∂D of PL are in one-to-one correspondence with the elements of the p th stable homotopy group $\pi_{n+p}(S^n)$ ($n \geq p + 1$) (\rightarrow 202 Homotopy Theory H).

Warning: The topological structure of an infinite-dimensional Hilbert or Banach space is quite different from that of a finite-dimensional Euclidean space. For instance, let X be a Hilbert space of infinite dimension, and let $D = \{x \in X \mid \|x\| < 1\}$ be its unit ball and $\partial D = \{x \in X \mid \|x\| = 1\}$ be the unit sphere of X . S. Kakutani (*Proc. Imp. Acad. Tokyo*, 19 (1943)) gave a fixed-point free continuous mapping of \bar{D} into itself if X is separable [1–3]. Thus naive generalization of the Brouwer fixed-point theorem is no longer true in infinite-dimensional spaces. V. Klee and C. Bessaga [17] proved that the unit sphere ∂D is C^∞ -diffeomorphic to X for an arbitrary Hilbert space X of infinite dimension. N. H. Kuiper proved that the group of invertible continuous linear operators on X is contractible if X is separable [3]. All this is in striking contrast to the well-known facts for finite-dimensional spaces (\rightarrow 202 Homotopy Theory, 427 Topology of Lie Groups and Homogeneous Spaces). This is the reason why compactness assumptions are made in the theorems mentioned above.

E. Calculus in Banach (or Locally Convex) Spaces

When one considers a nonlinear operator, it often happens that the domain and the range are neither linear spaces nor their open subsets. The domain might be a space of all smooth mappings of a compact manifold into another, and so might be the range. Such spaces have no linear structure, and hence linearity or semilinearity do not make sense in

general. The concept of infinite-dimensional manifolds is therefore introduced of necessity in nonlinear functional analysis.

Definition of differentiable mappings. Let E and F be real Banach spaces and let $L(E, F)$ ($= L(E)$ if $E = F$) be the Banach space of all bounded linear operators with uniform operator norm. Let U be an open subset of E and x a point of U . A mapping (= nonlinear operator) f of U into F is called **Gâteaux differentiable** at x if $\lim_{t \rightarrow 0} t^{-1}(f(x+ty) - f(x)) = df(x, y)$ exists for any $y \in E$. $df(x, y)$ is called the **Gâteaux derivative** of f at x . f is called **Fréchet differentiable** at x if there exists a linear operator $A \in L(E, F)$ such that $\lim_{y \rightarrow 0} \|f(x+y) - f(x) - Ay\|/\|y\| = 0$. A is called the **Fréchet derivative** of f at x and is denoted by $df(x)$ or $f'(x)$. f is Fréchet differentiable in U if and only if it is Gateaux differentiable, $df(x, y)$ is linear in y , and $\sup_{y \neq 0} \|df(x, y)\|/\|y\|$ is bounded [10].

Let U be an open subset of E . A mapping f of U into F is said to be of **class C^0** if it is continuous and to be of **class C^1** if it is Fréchet-differentiable at each point $x \in U$ and the differential $df(x) \in L(E, F)$ is continuous as a mapping of U into $L(E, F)$. The differential $df(x)$ is also called the **linearized operator**. If the mapping $df: U \rightarrow L(E, F)$ is of class C^{r-1} , then f is said to be of **class C^r** . $d(d^{r-1}f)(x)$ is written as $d^r f(x)$, and called the **r th differential** at x . $d^r f(x)$ is an r -linear, bounded, symmetric operator of $E \times \dots \times E$ (r times) into F . f is said to be of **class C^∞** if f is of class C^r for every r . For an open subset V of F , $f: U \rightarrow V$ is called a **C^r diffeomorphism** if f is a bijection and both f and f^{-1} are of class C^r .

A C^r mapping f of U into F is called a **Fredholm mapping (Fredholm map)** if $df(x) \in L(E, F)$ is a linear \dagger Fredholm operator for every $x \in U$. Since $\text{Ind } df(x)$ is constant if U is connected, that integer $\text{Ind } df(x)$ is called the **index** of f .

F. Taylor's Theorem and Its Converse

Let $f: U \rightarrow F$ be of class C^r ($r \geq 1$). A generalized Taylor theorem claims that f can be approximated by a polynomial mapping: Let $x \in U$ and $y \in E$ be sufficiently close to 0 so that $x + ty \in U$ for $0 \leq t \leq 1$. Then

$$f(x+y) = \sum_{k=0}^r \frac{1}{k!} d^k f(x)(y, \dots, y) + \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} \times \{d^r f(x+ty) - d^r f(x)\}(y, \dots, y) dt.$$

Let $L_{\text{sym}}^k(E, F)$ be the Banach space of all k -linear, bounded, symmetric operators of $E \times \dots \times E$ into F with the uniform topology. If for every k , $0 \leq k \leq r$, there exists a continuous

mapping $\varphi_k: U \rightarrow L_{\text{sym}}^k(E, F)$ such that

$$f(x+y) = \sum_{k=0}^r \frac{1}{k!} \varphi_k(x)(y, \dots, y) + o(\|y\|^r)$$

at every $x \in U$ and y sufficiently close to 0, then $f: U \rightarrow F$ is of class C^r , and $d^k f(x) = \varphi_k(x)$.

G. The Implicit Function Theorem

Using the notation above, let $f: U \rightarrow V$ be of class C^r , $r \geq 1$, and assume that $0 \in U$, $0 \in V$, and $f(0) = 0$, where 0 is the origin of E or F . Suppose there is an $A \in L(F, E)$, called the **right inverse** of $df(0)$, such that $df(0)A = 1_F$ (the identity). Then the following assertions hold: (i) The image of F under A , AF , is a closed subspace of E , and $E = \text{Ker } df(0) \oplus AF$. (ii) There are neighborhoods U_1, U_2, V' of the zeros of $\text{Ker } df(0)$, AF , F , respectively, such that $U_1 \oplus U_2 \subset U$ and such that the mapping $g: U_1 \oplus U_2 \rightarrow U_1 \oplus V'$ defined by $g(u, v) = (u, f(u, v))$ is a C^r -diffeomorphism. Therefore, denoting the inverse of g by $h = (h_1, h_2)$, we have $h_1(u, w) \equiv u$ and $f(u, h_2(u, w)) \equiv w$. The latter means that the nonlinear equation $f(u, v) = w$ can be solved with respect to v .

H. Existence and Uniqueness of Integral Curves

Using the notation above, let f be a C^r mapping ($r \geq 1$) of U into E . Since $U \times E$ is the \dagger tangent bundle of U , $(x, f(x))$ can be regarded as a C^r tangent vector field on U . The equation of \dagger integral curves is $(d/dt)x(t) = f(x(t))$. A local existence and uniqueness of solutions is stated as follows: For an arbitrarily fixed $x \in U$, there are $\varepsilon > 0$ and an open neighborhood W of x such that there exists uniquely a C^r mapping h of $W \times (-\varepsilon, \varepsilon)$ into U satisfying $(d/dt)h(w, t) = f(h(w, t))$ and $h(w, 0) \equiv w$.

Using this fact, one can prove the **Frobenius theorem**: Let E' be a closed linear subspace of E with a direct summand E'' , and let $f: U \rightarrow L(E', E'')$ be a C^r mapping ($r \geq 1$) such that $f(0) = 0$. To each $x \in U$ one associates a closed linear subspace $D_x = \{v, f(x)v\} | v \in E'\}$. The disjoint union $D = \bigcup_{x \in U} D_x$ can be regarded as a subbundle of the tangent bundle $U \times E$. A mapping \tilde{u} of U into E is called a **cross section** of D if $\tilde{u}(x) \in D_x$ for every $x \in U$. D is called **involutive** if for any two C^1 cross sections \tilde{u}, \tilde{v} of D , the Lie bracket product $[\tilde{u}, \tilde{v}]$ defined by $[\tilde{u}, \tilde{v}](x) = d\tilde{u}(x)(\tilde{v}(x)) - d\tilde{v}(x)(\tilde{u}(x))$ is again a cross section of D . Now suppose D is an involutive subbundle of $U \times E$. Then for an arbitrarily fixed $x \in U$, there are a neighborhood W

of x and a C^r diffeomorphism h of W onto a neighborhood of 0 of E such that $dh(x)(D_x) \equiv E'$. This fact shows that an involutive subbundle can be trivialized by a suitable change of local coordinate systems.

I. Local Theories on Locally Convex Spaces

All local theories mentioned in Sections E–H are constructed on Banach spaces. However, it is important in concrete applications to construct these theories on a wider class of locally convex topological linear spaces.

Let E, F be \dagger locally convex topological linear spaces, and let U be an open subset of E . A mapping $f: U \rightarrow F$ is said to be of class C^0 if it is continuous. f is said to be of class C^r if f is of class C^{r-1} and the following is fulfilled: For every $x \in U$, there is an r -linear continuous symmetric mapping $d^r f(x)$ of $E \times \dots \times E$ into F such that $d^r f: U \times E \times \dots \times E \rightarrow F$ is continuous. If we put

$$F(y) = f(x+y) - f(x) - df(x)(y) - \dots - \frac{1}{r!} d^r f(x)(y, \dots, y)$$

for every y sufficiently close to $0 \in E$, then the mapping G defined by

$$G(t, y) = \begin{cases} F(ty)/t^r, & t \neq 0, \\ 0, & t = 0, \end{cases} \quad (t \in \mathbf{R})$$

is continuous at $(0, 0) \in \mathbf{R} \times E$. The definitions of C^∞ mappings and C^r diffeomorphisms are given as in Section E.

J. Implicit Function Theorems in Locally Convex Spaces

The implicit function theorem does not hold in general locally convex spaces. However, since it is useful for nonlinear problems, several sufficient conditions are presently being studied. The following are some of them. We assume that $f: U \rightarrow F$ is a C^r mapping ($r \geq 1$) such that $f(0) = 0$ and that $df(0): E \rightarrow F$ has a continuous right inverse.

(i) If $\dim F < \infty$, the implicit function theorem holds just as in Section G.

(ii) An implicit function theorem that can be applied to nonlinear elliptic differential operators can be restated as follows. Suppose E, F are \dagger projective limit of families of Banach spaces $\{E^k, k \geq d\}$, $\{F^k, k \geq d\}$ and $U = E \cap U^d$, where U^d is an open subset of E^d . If $f: U \rightarrow F$ can be extended to a C^r mapping ($r \geq 2$) of $E^k \cap U^d$ into F^k for every $k \geq d$, if f satisfies the following inequalities, called a linear estimate

with respect to $|\cdot|_k$:

$$\begin{aligned} |df(x+y)(v)|_k &\leq C(|y|_k|v|_d + |y|_d|v|_k) \\ &\quad + P_k(|y|_{k-1})|v|_{k-1}, \quad k > d, \\ |d^2 f(x+y)(u, v)|_k &\leq C(|y|_k|u|_d|v|_d + |y|_d|u|_k|v|_d \\ &\quad + |y|_d|u|_d|v|_k) + P_k(|y|_{k-1})|u|_{k-1}|v|_{k-1}, \quad k > d, \end{aligned} \quad (4)$$

where $|\cdot|_k$ is the norm in E^k or F^k , C is a positive constant independent of k , and P_k is a polynomial with positive coefficients, and if a right inverse A of $df(0)$ satisfies the inequality of Gårding type,

$$|Au|_k \leq C'|u|_k + D_k|u|_{k-1}, \quad k > d, \quad (5)$$

where $C' > 0$ independent of k and $D_k > 0$, then the implicit function theorem holds just as in Section G, and the obtained mapping h satisfies the same inequalities as (4) [11].

(iii) Nash-Moser implicit function theorem.

Though linear estimates such as (4) hold for many differential operators, the second inequality (5) is sometimes out of order, especially if f is a nonlinear hyperbolic operator. However, one can often obtain instead of (5) a weaker inequality:

$$|Au|_k \leq C'|u|_{k+s} + D_k|u|_{k+s-1}, \quad s > 0.$$

J. Nash [37] and J. Moser [12] approximated such an operator A by some smoothing operators and proved an implicit function theorem under a certain additional condition. The Nash-Moser implicit function theorem was successfully applied to many difficult problems, e.g., the embedding problem of Riemannian manifolds [37], the small divisor problem of celestial mechanics [36], free boundary problems (e.g., L. Hörmander *Arch. Rational Mech. Anal.*, 62 (1976), 1–52), and other problems (e.g., S. Klainerman, *Comm. Pure Appl. Math.*, 33 (1980), 43–101; M. Kuranishi, *Amer. Math. Soc. Proc. Symposia in Pure Math.*, 30 (1977), 97–105).

(iv) Analytic implicit function theorem.

In cases (ii) and (iii), the spaces E, F were given as projective limits of Banach spaces. On the contrary, H. Jacobowitz considered the case where E, F are \dagger inductive limits of Banach spaces. For instance, the space E of the smooth functions can be approximated by a family of Banach spaces $\{E_\varepsilon | \varepsilon > 0\}$ of all real analytic functions with ε as the radius of convergence. Under this circumstance, certain conditions for f and the right inverse of $df(0)$ yield an implicit function theorem [13].

(v) Mather's implicit function theorem [14]. The difficulty of implicit function theorems in Fréchet spaces is concentrated in the following fact: Even if $df(0)$ has a right inverse and even if x is sufficiently close to 0, $df(x)$ may not have

a right inverse. If the implicit function theorem holds, such a phenomenon should not happen. In cases (ii)–(iv), several functional-analytic conditions exclude this pathological phenomenon. In some special cases, these conditions can be replaced by algebraic ones. J. Mather, using his division theorem, proved an implicit function theorem that was applied to the theory of singularities.

K. Infinite-Dimensional Manifolds

A Hausdorff space M is called a C^r **Banach manifold** modeled on a Banach space E if the following conditions are satisfied: (a) M is covered by a family of open subsets $\{U_\alpha\}_{\alpha \in A}$. For each U_α there are an open subset V_α of E and a homeomorphism ψ_α of U_α onto V_α . Such a pair is called a local coordinate system or a local chart of M . (b) If $U_\alpha \cap U_\beta \neq \emptyset$, $\psi_\alpha \psi_\beta^{-1}$ is a C^r -diffeomorphism of $\psi_\beta(U_\alpha \cap U_\beta)$ onto $\psi_\alpha(U_\alpha \cap U_\beta)$. (c) The index set A is maximal among those that satisfy (a) and (b).

If M (resp. M') is a C^r Banach manifold modeled on E (resp. E'), then so is $M \times M'$ modeled on $E \oplus E'$. A mapping $f: M \rightarrow M'$ is said to be of class C^r if f expressed through local coordinate systems is of class C^r .

Suppose M is covered by a family of local charts $\{(U_\alpha, \psi_\alpha)\}_{\alpha \in A}$. On the disjoint union $\bigcup_{\alpha \in A} U_\alpha \times E$, define an equivalence relation \sim as follows: For $(x, u) \in U_\alpha \times E$, $(y, v) \in U_\beta \times E$, $(x, u) \sim (y, v)$ if and only if $x = y$ and $d(\psi_\alpha \psi_\beta^{-1})(x)v = u$. The set of equivalence classes is called the **tangent bundle** of M , which will be denoted by T_M . There is another definition of the tangent bundle which uses the ring of C^r functions on M . However, since E is not reflexive in general, the latter gives us a different vector bundle. T_M is a C^{r-1} Banach manifold modeled on $E \oplus E$. The correspondence which sends $(x, u) \in \bigcup_{\alpha \in A} U_\alpha \times E$ to x induces a C^{r-1} mapping of T_M onto M , called the **projection** of T_M , denoted by π .

A topological group G is called a **Banach-Lie group** if G is a C^∞ Banach manifold and the group operation $(g, h) \rightarrow g^{-1}h$ is a C^∞ mapping of $G \times G$ into G . If E is a Banach space, then $GL(E)$, the group of all invertible bounded linear operators, is a Banach-Lie group under the uniform topology. If E is a Hilbert space, then the group of all unitary operators is also a Banach-Lie group under the same topology.

The concepts of manifolds and Lie groups are similarly defined when the model space is a Hilbert space or a Fréchet space. These are called a **Hilbert manifold** and a **Fréchet manifold**, respectively. For some differential

topologies on separable Hilbert manifolds \rightarrow
279 Morse Theory E.

L. Structures on Infinite-Dimensional Manifolds

Suppose M is a C^{r+1} Hilbert manifold modeled on E . At each $x \in M$, the tangent space $T_x M = \pi^{-1}(x)$ is a Hilbert space, linear-homeomorphic to E . M is called a C^r **Riemannian manifold** if there is defined an inner product $\langle u, v \rangle_x$ on each $T_x M$ such that $\langle \cdot, \cdot \rangle_x$ is of class C^r with respect to x . Existence of such a structure is ensured by using a partition of unity if M is paracompact.

Let M be a C^r Banach manifold ($r \geq 1$) modeled on a Banach space E . Each tangent space $T_x M$ is linear homeomorphic to E . M is called a C^r **Finsler manifold** if there is a norm $|u|_x$ defined on each $T_x M$ such that $| \cdot |_x$ is continuous with respect to x . A paracompact C^1 Banach manifold can have a C^1 Finsler structure.

Let M be a C^{r+2} Fréchet manifold ($r \geq 0$) and $C^{r+1}(M)$, $\Gamma^{r+1}(T_M)$ the spaces of all C^{r+1} functions and of all C^{r+1} vector fields on M , respectively. A bilinear mapping ∇ of $\Gamma^{r+1}(T_M) \times \Gamma^{r+1}(T_M)$ into $\Gamma^r(T_M)$ is called an **affine connection** on M if ∇ satisfies $\nabla_{f\tilde{u}}\tilde{v} = f\nabla_{\tilde{u}}\tilde{v}$, $\nabla_{\tilde{u}}f\tilde{v} = (\tilde{u}f)\tilde{v} + f\nabla_{\tilde{u}}\tilde{v}$ for every $\tilde{u}, \tilde{v} \in \Gamma^{r+1}(T_M)$, $f \in C^{r+1}(M)$. For an affine connection ∇ , $T(\tilde{u}, \tilde{v}) = \nabla_{\tilde{u}}\tilde{v} - \nabla_{\tilde{v}}\tilde{u} - [\tilde{u}, \tilde{v}]$ is called the **torsion tensor**, and $R(\tilde{u}, \tilde{v}) = \nabla_{\tilde{u}}\nabla_{\tilde{v}} - \nabla_{\tilde{v}}\nabla_{\tilde{u}} - \nabla_{[\tilde{u}, \tilde{v}]}$ is called the **curvature tensor** of ∇ . If M is a Riemannian manifold, then there exists a unique affine connection without torsion which leaves the Riemannian inner product parallel.

M. Local Linearization Theorems

Let M be a C^{r+1} Banach manifold ($r \geq 1$) modeled on E . Let \tilde{u} be a C^r vector field on M such that $\tilde{u}(x) \neq 0$ at $x \in M$. Then there are a neighborhood U_x of x and a C^r diffeomorphism ψ of U_x onto an open subset V of E such that $d\psi\tilde{u}(\psi^{-1}(y)) \equiv (y, v)$, $v \in E$, for every $y \in V$, where v does not depend on y .

N. Morse Lemma

Let M be a C^{r+2} Hilbert manifold and f be an \mathbf{R} -valued C^{r+2} function. Suppose x is a critical point of f , i.e., $df(x) = 0$. x is called a **nondegenerate critical point** if $d^2f(x)$ is a nondegenerate bilinear form. For such x there are a neighborhood U_x of x and a C^r diffeomorphism ψ of U_x onto an open neighborhood of 0 of the model space E such that $\psi(x) = 0$, and $f(\psi^{-1}(y)) = |Py|^2 - |(1-P)y|^2$, where P is an orthogonal

projection in E . $i_x = \dim(1 - P)E$ ($0 \leq i_x \leq \infty$) is called the **index** of the critical point x of f .

O. Submanifolds

A subset N of a C^r Banach manifold M modeled on E is called a C^r **submanifold** if at each point $x \in N$ there are a neighborhood U_x of x and a C^r -diffeomorphism ψ of U_x onto an open neighborhood V of 0 of E such that $\psi(x) = 0$ and $\psi(U_x \cap N) = V \cap F$, where F is a closed linear subspace of E . There are some other definitions of submanifolds. One of them requires in addition that F be a direct summand of E , and another uses instead of $U_x \cap N$ its connected component containing x . In the latter definition, a submanifold is not necessarily locally closed.

P. Sard-Smale Theorem

Although it is not easy to define nontrivial measures on infinite-dimensional manifolds, the concept "almost everywhere" can sometimes be replaced by that of residual sets. A subset of a topological space is called a \dagger residual set if it contains an intersection of countably many open dense subsets. A residual set in a complete metric space or in a \dagger Baire space is dense. S. Smale [15] extended \dagger Sard's theorem to infinite-dimensional manifolds as follows: Let M, N be C^r Banach manifolds and $f: M \rightarrow N$ be a C^r Fredholm mapping. If M is separable and $r > \max\{0, \text{Ind}(df(x))\}$ for each $x \in M$, then $R_f = N - f(C)$ is a residual subset of N , where C is the set of critical points of f , i.e., a point x where $df(x)$ is not surjective.

Q. Calculus of Variations and Infinite-Dimensional Manifolds

Many problems in the calculus of variations can be understood as problems seeking critical points of functions defined on infinite-dimensional manifolds. R. Palais and S. Smale set up the following Condition C and fixed a category of functions where the critical points can be chased through gradient-like vector fields [6].

Palais-Smale Condition C. Let f be a C^1 function on a C^1 Finsler manifold M . If S is any subset of M on which f is bounded but on which $|df(x)|$ is not bounded away from 0 , then there is a critical point of f adherent to S .

In general, it is not easy to examine Condition C for a concrete f . However, many concrete problems, where the Euler equations are nonlinear elliptic, satisfy Condition C.

Morse theory. Let M be a C^∞ complete Riemannian manifold and f a C^∞ function bounded below satisfying Condition C and having only nondegenerate critical points. Then, using the Morse lemma, one can make a \dagger handlebody decomposition of M by the same method as in the case of finite-dimensional manifolds (\rightarrow 279 Morse Theory).

Lyusternik-Shnirel'man theory. This theory, constructed on finite-dimensional manifolds, can be extended naturally to Finsler manifolds. Let M be a complete C^2 Finsler manifold and f a C^2 function satisfying Condition C and bounded below. Then f has at least $\text{cat}(M)$ critical points, where $\text{cat}(M) = m$ means that M can be covered by m closed contractible subsets of M but not by $m - 1$ ones. If there is no such integer, then we set $\text{cat}(M) = \infty$.

Both Morse theory and Lyusternik-Shnirel'man theory have been successfully employed in the global theory of the calculus of variations.

R. Bifurcation Theory

Bifurcation theory concerns itself with the structure of the zeros of the functional equation of w with a parameter λ :

$$G(\lambda, w) = 0. \quad (6)$$

In general, the state w satisfying (6) represents the equilibrium (time-independent or stationary) solution of the evolution equation

$$w_t = G(\lambda, w) \quad \text{and} \quad w(0) = w_0. \quad (7)$$

Here the evolution equation itself stems from a mathematical model describing natural phenomena, $w = w(t)$ stands for the state at time t , and λ is the set of parameters representing the physical environment. For example, in the \dagger Navier-Stokes equation appearing in fluid dynamics, $w(t)$ represents the unknown velocity field at time t and λ is the \dagger Reynolds number. It is important to study bifurcation phenomena because they typically accompany the transition to instability of the state when some characteristic parameter passes through a certain value, called a **critical value**.

Let X, Y , and Λ be real Banach spaces, and let $G(\lambda, w)$ be a mapping from $\Lambda \times X$ to Y . Suppose that there exists a mapping $\tilde{w}(\lambda): \Lambda \rightarrow X$ satisfying $G(\lambda, \tilde{w}(\lambda)) = 0$. One calls $(\lambda, \tilde{w}(\lambda))$ a trivial solution of (6).

$(\lambda_0, \tilde{w}(\lambda_0))$ is called a **bifurcation point** of $G(\lambda, w)$ (with respect to the trivial solution) if in any neighborhood of $(\lambda_0, \tilde{w}(\lambda_0))$ there exists a nontrivial solution of (6). (In general, there may appear another type of solution that is not connected with the trivial solution [20]).

Assume that $G(\lambda, w)$ is of class C^1 in some neighborhood of $(\lambda_0, \tilde{w}(\lambda_0))$ in $\Lambda \times X$. Then it follows from the implicit function theorem that $(\lambda_0, \tilde{w}(\lambda_0))$ is not a bifurcation point if $G_w(\lambda_0, \tilde{w}(\lambda_0))$, the Fréchet derivative of G with respect to w , is nonsingular.

S. The Principle of Linearized Stability

Closely tied to the phenomenon of bifurcation is the property of **stability**. Suppose that the dynamics of a physical system are governed by (7). Let $w(t; w_0)$ denote the solution of (7). An equilibrium solution $\tilde{w} = \tilde{w}(\lambda)$ is called **stable** if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|w(t; w_0) - \tilde{w}\| < \varepsilon$ for all $t > 0$ whenever $\|w_0 - \tilde{w}\| < \delta$. Furthermore, \tilde{w} is said to be **asymptotically stable** if, in addition, $w(t; w_0) \rightarrow \tilde{w}$ as $t \rightarrow \infty$.

By the **principle of linearized stability** we mean that the stability of an equilibrium solution \tilde{w} is determined formally by the spectrum of the linearized operator $G_w(\lambda, \tilde{w}(\lambda))$. Assume that $G_w(\lambda, \tilde{w}(\lambda))$ has only a point spectrum. Then (i) if $\tilde{w}(\lambda)$ is stable, the spectrum of $G_w(\lambda, \tilde{w}(\lambda))$ is contained in $\{z \in \mathbb{C} \mid \operatorname{Re} z \leq 0\}$, and (ii) if the spectrum of $G_w(\lambda, \tilde{w}(\lambda))$ is contained in $\{z \in \mathbb{C} \mid \operatorname{Re} z < 0\}$, \tilde{w} is asymptotically stable.

Suppose that as λ crosses a certain value λ_0 , one or more eigenvalues of $G_w(\lambda, \tilde{w}(\lambda))$ cross the imaginary axis from the left to the right half-plane, where $\tilde{w}(\lambda)$ is the known equilibrium solution. This is precisely the situation when $\tilde{w}(\lambda)$ becomes unstable. For notational simplicity, we put $u = w - \tilde{w}(\lambda)$ and $F(\lambda, u) = G(\lambda, u + \tilde{w}(\lambda))$.

T. Bifurcation from Simple Eigenvalues

Take $\Lambda = \mathbb{R}$. Let $F \in C^2(\mathbb{R} \times X, Y)$ be such that $F(\lambda, 0) = 0$ for any real λ . Set $L_0 = F_u(\lambda_0, 0)$, $L_1 = F_{\lambda u}(\lambda_0, 0)$, and suppose that (i) $\operatorname{Ker}(L_0)$ is spanned by $u_0 \neq 0$; (ii) $\operatorname{codim} R(L_0) = 1$; and (iii) $L_1 u_0 \notin R(L_0)$, where $R(L_0)$ denotes the range of L_0 . Then there exists a C^1 -curve $(\lambda, \psi): (-\delta, \delta) \rightarrow \mathbb{R} \times X$ defined on some interval $(-\delta, \delta)$ such that $\lambda(0) = \lambda_0$, $\psi(0) = 0$, and $F(\lambda(s), s(u_0 + \psi(s))) = 0$ for any $s \in (-\delta, \delta)$. Moreover, in a neighborhood of $(\lambda_0, 0)$ any zero of F either lies on this curve or is a trivial solution (M. G. Crandall and P. H. Rabinowitz, *J. Functional Anal.*, 8 (1971)).

In the above case, there exist only three possible situations of the curves (λ, φ) and $(\lambda, 0)$, called subcritical, supercritical, and transcritical bifurcations. In the third case, there occurs the so-called **exchange of stability** [19–21].

U. Bifurcation of Periodic Solutions (Hopf Bifurcation Theorem)

If $u(\lambda)$ loses stability by virtue of a pair of complex conjugate eigenvalues crossing the imaginary axis, then under suitable conditions one can prove the existence of bifurcating time-periodic solutions of (7). Rewrite equation (7) as

$$u_t + L_0 u + g(\lambda, u) = 0. \quad (8)$$

Suppose (i) $L_0: D(L_0) \subset X \rightarrow X$ is a densely defined linear operator on X such that $-L_0$ generates a strongly continuous semigroup on X , which is holomorphic on $X^{\mathbb{C}}$ (= the complexification of X). L_0 has compact resolvent; i is a simple eigenvalue of L_0 , and $ni \notin \sigma(L_0)$, the spectrum of L_0 , for $n = 0, 2, 3, \dots$. As a consequence of (i), if $r > -\operatorname{Re} \lambda$ for any $\lambda \in \sigma(L_0)$, then the fractional power $(rI + L_0)^\alpha$ for $\alpha \geq 0$ is well defined. Because their domains are independent of r , one can set $X_\alpha = D((rI + L_0)^\alpha)$, which are Banach spaces under the norm $\|u\|_\alpha = \|(rI + L_0)^\alpha u\|$. Suppose (ii) there exist an $\alpha \in [0, 1)$ and a neighborhood \mathcal{O} of $(0, 0)$ in $\mathbb{R} \times X$ such that $g \in C^2(\mathcal{O}, X_\alpha)$, where $C^k(\mathcal{O}, X_\alpha)$ denotes the space of all X_α -valued C^k functions defined on \mathcal{O} . Moreover $g(\lambda, 0) = 0$ if $(\lambda, 0) \in \mathcal{O}$, and $g_u(0, 0) = 0$. (iii) Let $\beta = \beta(\lambda)$ be a continuously differentiable function defined in a neighborhood of 0 such that $\beta(\lambda) \in \sigma(L_0 + g_u(\lambda, 0))$ and $\beta(0) = i$. Suppose that $\operatorname{Re} \beta'(0) \neq 0$. If assumptions (i), (ii) and (iii) are satisfied, then there exist a positive δ and continuously differentiable functions $(\rho, \lambda, u): (-\delta, \delta) \rightarrow \mathbb{R}^2 \times C^0(\mathbb{R}, X_\alpha)$ such that (a) for $0 < |s| < \delta$, $u(s)$ is a $2\pi\rho(s)$ -periodic solution of period $2\pi\rho(s)$ of (8) corresponding to $\lambda = \lambda(s)$; (b) $\rho(0) = 1$, $\lambda(0) = 0$, $u(0) = 0$, and $u(s) \neq 0$ if $s \neq 0$; and (c) any $2\pi\rho$ -periodic solution of (8) in $C_{2\pi\rho}^0(\mathbb{R}, X_\alpha)$ (= the space of $2\pi\rho$ -periodic continuous functions with value in X_α) with $|\rho - 1|$, $|\lambda|$ and $\|u\|$ sufficiently small is of the above form for some $|s| < \delta$ up to a translation of the real line. Moreover, if $g \in C^{k+1}(\mathcal{O}, X_\alpha)$, then the functions ρ, λ, u are of class C^k .

V. The Lyapunov-Schmidt Procedure

Suppose that $L(\lambda) = F_u(\lambda, 0)$ has at $\lambda = 0$ an n -fold eigenvalue at the origin, i.e., $\dim \operatorname{Ker} L(0) = n$, and assume that $X \subset Y$. Let $\operatorname{Ker} L(0)$ be spanned by $\varphi_1, \dots, \varphi_n$, and let P be the projection onto this linear subspace that commutes with $L(0)$. Then P must take the form $Pu = \sum_{j=1}^n \langle u, \varphi_j^* \rangle \varphi_j$, where the $\varphi_j^* \in Y^* \subset X^*$ are null functions of the adjoint operator $L(0)^*$ and $\langle \varphi_j, \varphi_k^* \rangle = \delta_{jk}$. P is a linear operator from Y to Y ; hence it can be regarded as a mapping from X to X as well, and $Q = I - P$ is a projec-

tion onto the range of $L(0)$ in Y . Using P and Q , the equation $F(\lambda, u) = 0$ can be decomposed into the system of equations

$$QF(\lambda, v + \psi) = 0 \quad \text{and} \quad PF(\lambda, v + \psi) = 0, \quad (9)$$

where $v = Pu$ and $\psi = Qu$. One solves the first equation for $\psi = \psi(\lambda, v)$ by the implicit function theorem, and then, substituting this into the second, one has the **bifurcation equation**

$$\tilde{F}(\lambda, v) = PF(\lambda, v + \psi(\lambda, v)) = 0. \quad (10)$$

Solutions of the bifurcation equation are in one-to-one correspondence with solutions of the original system sufficiently close to the bifurcation point.

There exists another method of reducing an infinite-dimensional problem to a finite-dimensional one, called the **center manifold theorem** [21, 39].

W. A Global Result

The following global result is due to P. H. Rabinowitz (*J. Functional Anal.*, 7 (1971)). Assume that $F(\lambda, u) = u - \lambda Lu + H(\lambda, u)$, where L is a compact linear operator and $H: \mathbf{R} \times X \rightarrow X$ is a compact mapping with $H(\lambda, u) = o(\|u\|)$ at 0 uniformly on bounded λ -intervals. Then, if μ^{-1} is an eigenvalue of L of odd multiplicity, $(\mu, 0)$ is a bifurcation point for F with respect to the trivial solution. Moreover, the closure of the set of nontrivial zeros of F contains a component that meets $(\mu, 0)$ and either is unbounded in $\mathbf{R} \times X$ or meets $(\hat{\mu}, 0)$, where $\mu \neq \hat{\mu}$ and $\hat{\mu}^{-1}$ is an eigenvalue of L .

The beginning of bifurcation theory seems to be in the celebrated work of H. Poincaré (*Acta Math.*, 7 (1885)).

X. Abstract Cauchy Problems

Suppose that we are given an **abstract Cauchy problem**

$$\frac{du}{dt} = Au \quad (t > 0), \quad (11)$$

$$u(+0) = a \quad (12)$$

in a Banach space X , where $a \in X$ and the nonlinear operator A is assumed, for simplicity, to be independent of t . If the domain $D(A)$ of A coincides with X and A is † Lipschitz continuous, we can reduce the abstract Cauchy problem to the † integral equation of Volterra type

$$u(t) = a + \int_0^t Au(s) ds,$$

and by applying the iteration procedure we can easily show that the abstract Cauchy problem has a uniquely determined solution.

A similar treatment works for a singular but mildly nonlinear A of the form $A = L + N$ if the linear operator L is the generator of a † strongly continuous semigroup e^{tL} in the sense of Hille-Yosida theory (\rightarrow 378 Semigroups of Operators and Evolution Equations) and the nonlinear mapping N is Lipschitz continuous. In this case, we reduce the problem to the integral equation

$$u(t) = e^{tL}a + \int_0^t e^{(t-s)L}Nu(s) ds.$$

Here, if we merely have to assure ourselves of the local existence of the solution, then N can merely be locally Lipschitz continuous; and in this case, N can even be singular to some extent if e^{tL} is † holomorphic in t [22, 23]. A typical application of this procedure was made by P. Sobolevskii, T. Kato, and H. Fujita to the Navier-Stokes equation (\rightarrow 204 Hydrodynamical Equations; 205 Hydrodynamics) to construct regular solutions [22].

† Galerkin's method (\rightarrow 304 Numerical Solution of Partial Differential Equations) is sometimes quite convenient for obtaining († weak) solutions of (11) and (12) [5]. Again, typical applications were made to the Navier-Stokes equations by E. Hopf and others [5, 24, 25]. To prove the convergence of approximate solutions constructed by Galerkin's method and subject to so-called energy estimates, we often make use of Aubin's compactness theorem concerning vector-valued functions [5].

Remarkable developments have taken place since 1964 for the case where A is dissipative, i.e., $-A$ is accretive. F. Browder [26] proved that if X is a Hilbert space and A is dissipative, then the mere continuity of A is sufficient for the † well-posedness of (11) and (12). Then Y. Kōmura [27] brought about a crucial advance by showing that if A is a (possibly multivalued and) maximal dissipative operator with $D(A)$ dense in the Hilbert space X , then (11) and (12) are uniquely solvable for any $a \in D(A)$. Furthermore, he founded the theory of **nonlinear semigroups of operators** by establishing a † nonlinear version of the Hille-Yosida theory for semigroups of † nonexpansive operators in Hilbert spaces [28]. Subsequent developments and applications were made in various directions by T. Kato (*J. Math. Soc. Japan*, 19 (1967)), M. Crandall and A. Pazy (*J. Functional Anal.*, 3 (1969)), H. Brézis and A. Pazy (*J. Functional Anal.*, 6 (1970)), S. Oharu (*J. Math. Soc. Japan*, 22 (1970)), M. Crandall and T. Liggett [29], Y. Konishi (*Proc. Japan Acad.*, 47 (1971)), and others. While Kōmura's original proof was based on the **Yosida approximation** $A(I - \varepsilon A)^{-1}$ ($\varepsilon \rightarrow 0$) for A , it is proved in [29] that in any Banach space a dissipative operator A that is maxi-

mal in a certain sense generates a nonlinear semigroup T_t by the **exponential formula**

$$T_t x = \lim_{n \rightarrow \infty} \left(I - \frac{t}{n} A \right)^{-n} x. \quad (13)$$

The scope of applications of the generating theorem (13) can be seen, e.g., in B. K. Quinn (*Comm. Pure Appl. Math.*, 24 (1971)), M. G. Crandall (*Israel J. Math.*, 12 (1972)), Y. Konishi (*Proc. Japan Acad.*, 48 (1972) and *J. Math. Soc. Japan*, 25 (1973)), and S. Aizawa (*Hiroshima Math. J.*, 6 (1976)).

Y. Nonlinear Semigroups in Banach Lattices

Let X be a Banach lattice (\rightarrow 310 Ordered Linear Spaces). An operator $A: X \supset D(A) \rightarrow X$ is said to be **dispersive** if for all $x, y \in D(A)$,

$$\|(x - y)^+\| \leq \|(x - y - \lambda(Ax - Ay))^+\| \quad (\forall \lambda > 0).$$

When a dispersive operator A satisfies the range condition $R(I - \lambda A) \supset D(A)$ for any $\lambda > 0$, it generates an **order-preserving semigroup** $T_t = e^{tA}$ on $\overline{D(A)}$: $\|(T_t x - T_t y)^+\| \leq \|(x - y)^+\|$ for $t > 0$. We have therefore the preservation of order: $x \leq y$ implies $T_t x \leq T_t y$. We can prove in particular that the order of initial data is inherited by the solutions of a nonlinear heat equation (Y. Konishi).

Remark. Various pathological phenomena arise when we do not restrict the form of A in the abstract Cauchy problem (11), (12). We cite merely the "blowing up" of solutions of Cauchy problems for *nonlinear heat equations (\rightarrow 291 Nonlinear Problems) and the nonlinear wave propagations described by nonlinear *Schrödinger equations (J. B. Bailon, T. Cazenave, and M. Figueira).

Z. Abstract Cauchy-Kovalevskaya Theorem in a Scale of Banach Spaces

T. Yamanaka (*Comment. Math. Univ. St. Paul*, 9 (1960)), L. V. Obsyannikov (*Soviet Math. Doklady*, 6 (1965) and 12 (1971)), F. Trèves (*Trans. Amer. Math. Soc.*, 150 (1970)), L. Nirenberg (*J. Differential Geometry*, 6 (1972)), and T. Nishida [30] discussed abstract treatments of classical Cauchy-Kovalevskaya theorem for partial differential equations: Let $S = \{B_\rho\}_{\rho > 0}$ be a collection of Banach spaces depending on the real parameter $\rho > 0$. Let $\|u\|_\rho$ denote the norm of an element $u \in B_\rho$. The collection S is called a **scale of Banach spaces** if, for any ρ and $\rho' < \rho$, $B_\rho \subset B_{\rho'}$ and $\|u\|_{\rho'} \leq \|u\|_\rho$ for any u in B_ρ . Consider in S the initial value problem of the form

$$\frac{du}{dt} = F(u(t), t), \quad |t| < \delta, \quad \text{and} \quad u(0) = 0. \quad (14)$$

286 Ref.

Nonlinear Functional Analysis

Assume the following conditions on F : (i) For some numbers $R > 0$, $\eta > 0$, $\rho_0 > 0$ and every pair of numbers ρ, ρ' such that $0 < \rho' < \rho < \rho_0$, $(u, t) \rightarrow F(u, t)$ is a continuous mapping of $\{u \in B_\rho \mid \|u\|_\rho < R\} \times \{t \mid |t| < \eta\}$ into $B_{\rho'}$. (ii) For any $\rho' < \rho < \rho_0$ and all $u, v \in B_\rho$ with $\|u\|_\rho < R$, $\|v\|_\rho < R$, and for any t , $|t| < \eta$, F satisfies $\|F(u, t) - F(v, t)\|_{\rho'} < C\|u - v\|_\rho / (\rho - \rho')$, where C is a constant independent of t, u, v, ρ , or ρ' . (iii) $F(0, t)$ is a continuous function of t , $|t| < \eta$, with values in B_ρ for every $\rho < \rho_0$ and satisfies, with a fixed constant K , $\|F(0, t)\|_\rho \leq K/(\rho_0 - \rho)$, $0 \leq \rho < \rho_0$.

Abstract Cauchy-Kovalevskaya theorem.

Under the preceding hypotheses there is a positive constant M such that there exists a unique function $u(t)$ which for every positive $\rho < \rho_0$ and $|t| < M(\rho_0 - \rho)$ is a continuously differentiable function of t with values in B_ρ , $\|u(t)\|_\rho < R$, and satisfies (14) (\rightarrow also [31]).

These results cover a theorem of M. Nagumo (*Japan. J. Math.*, 18 (1948)), which generalizes the classical Cauchy-Kovalevskaya theorem.

References

- [1] L. Nirenberg, Topics in nonlinear functional analysis, Lecture notes, New York Univ., 1973–1974.
- [2] J. T. Schwartz, Nonlinear functional analysis, Gordon & Breach, 1969.
- [3] M. Berger, Nonlinearity and functional analysis, Academic Press, 1977.
- [4] G. J. Minty, Monotone (nonlinear) operators in Hilbert space, *Duke Math. J.*, 29 (1962), 341–346.
- [5] J. L. Lions, Quelques méthodes de résolution des problèmes aux limites nonlinéaires, Dunod, 1969.
- [6] J. K. Hale, Ordinary differential equations, Interscience, 1969.
- [7] M. A. Krasnosel'skii, Topological methods in the theory of nonlinear integral equations, Pergamon, 1964. (Original in Russian, 1956.)
- [8] A. Friedman, Partial differential equations of parabolic type, Prentice-Hall, 1964.
- [9] O. A. Ladyzhenskaya and N. N. Ural'tseva, Linear and quasilinear elliptic equations, Academic Press, 1968. (Original in Russian, 1956.)
- [10] E. Hille and R. S. Phillips, Functional analysis and semigroups, Amer. Math. Soc. Coll. Publ., 31 (1957).
- [11] H. Omori, Infinite-dimensional Lie transformation groups, Lecture notes in math. 427, Springer, 1974.
- [12] J. Moser, A new technique for construction of solutions of nonlinear differential equa-

- tions, Proc. Nat. Acad. Sci. US, 47 (1961), 1824–1831.
- [13] H. Jacobowitz, Implicit function theorems and isometric embeddings, *Ann. Math.*, 95 (1972), 191–225.
- [14] J. Mather, Stability of C^∞ mappings I, *Ann. Math.*, 87 (1968), 89–104.
- [15] S. Smale, An infinite-dimensional version of Sard's theorem, *Amer. J. Math.*, 87 (1965), 861–866.
- [16] R. Palais and S. Smale, A generalized Morse theory, *Bull. Amer. Math. Soc.*, 70 (1964), 165–172.
- [17] C. Bessaga, Every infinite-dimensional Hilbert space is diffeomorphic with its unit sphere, *Bull. Acad. Sci. Polon.*, 14 (1966), 27–31.
- [18] R. Palais, Lusternik-Schnirelman theory on Banach manifolds, *Topology*, 5 (1966), 115–132.
- [19] M. G. Crandall and P. H. Rabinowitz, *Mathematical theory of bifurcation, Bifurcation Phenomena in Mathematical Physics and Related Topics*, C. Bardos and D. Bessis (eds.), Reidel, 1980.
- [20] J. E. Marsden, Qualitative methods in bifurcation theory, *Bull. Amer. Math. Soc.*, 84 (1978), 1125–1148.
- [21] D. H. Sattinger, Bifurcation and symmetry breaking in applied mathematics, *Bull. Amer. Math. Soc.*, new ser., 3 (1980), 779–819.
- [22] T. Kato, Nonlinear evolution equations in Banach spaces, *Amer. Math. Soc. Proc. Symposia in Appl. Math.*, 17 (1965), 50–67.
- [23] P. E. Sobolevskii, Equations of parabolic type in a Banach space, *Amer. Math. Soc. Transl.*, (2) 49 (1966), 1–62. (Original in Russian, 1961.)
- [24] E. Hopf, Über die Anfangswertaufgabe für die hydrodynamischen Grundgleichungen, *Math. Nachr.*, 4 (1951), 213–231.
- [25] O. A. Ladyzhenskaya, *The mathematical theory of viscous incompressible flow*, Gordon & Breach, 1963. (Original in Russian, 1961).
- [26] F. Browder, Nonlinear equations of evolution, *Ann. Math.*, 80 (1964), 485–523.
- [27] Y. Kōmura, Nonlinear semigroups in Hilbert space, *J. Math. Soc. Japan*, 19 (1967), 493–507.
- [28] Y. Kōmura, Differentiability of nonlinear semigroups, *J. Math. Soc. Japan*, 21 (1969), 375–402.
- [29] M. G. Crandall and T. M. Liggett, Generation of semigroups of nonlinear transformations on general Banach spaces, *Amer. J. Math.*, 93 (1971), 265–298.
- [30] T. Nishida, A note on a theorem of Nirenberg, *J. Differential Geometry*, 12 (1977), 629–633.
- [31] T. Kano and T. Nishida, Sur les ondes de surface de l'eau avec une justification mathématique des équations des ondes en eau peu profonde, *J. Math. Kyoto Univ.*, 19 (1979), 335–370.
- [32] Yu. G. Borisovich, V. G. Zvyagin, and Yu. I. Saponov, Nonlinear Fredholm maps and the Leray-Schauder theory, *Russian Math. Surveys*, 32 (1977), 1–54. (Original in Russian, 1977.)
- [33] J. Leray and J. Schauder, Topologie et équation fonctionnelles, *Ann. Sci. Ecole Norm. Sup.*, 51 (1934), 45–78.
- [34] L. Lyusternik and L. Shnirel'man, *Méthode topologique dans les problèmes variationnels*, Herman, 1934. (Original in Russian, 1930.)
- [35] M. Morse, Calculus of variations in the large, *Amer. Math. Soc. Colloq. Publ.*, 18 (1934).
- [36] J. Moser, A rapidly convergent iteration method and nonlinear partial differential equations I, II, *Ann. Scuola Norm. Pisa*, 20 (1966), 265–315, 499–535.
- [37] J. Nash, The embedding problem for Riemannian manifolds, *Ann. Math.*, 63 (1965), 20–63.
- [38] K. D. Elworthy and A. J. Tromba, Differential structures and Fredholm maps, *Amer. Math. Soc. Proc. Symposia in Pure Math.*, 15 (1970), 45–94.
- [39] J. Carr, *Applications of centre manifold theory*, Springer, 1981.

287 (XX.33)

Nonlinear Lattice Dynamics

A. Lattice Dynamics

In order to elucidate certain characteristic features of nonlinear waves, **one-dimensional lattice** models have been studied. Around 1953, E. Fermi et al. performed computer experiments on nonlinear lattices to verify a generally accepted belief that nonlinear coupling between the normal modes of harmonic oscillators would lead to complete energy sharing between these modes. To their surprise, their nonlinear lattices yielded very little energy sharing at all; on the contrary, the interactions resulted in the recurrence of the initial state. These results were later interpreted in terms of solitons (\rightarrow 387 Solitons), i.e., nonlinear waves that preserve identity despite mutual interaction.

The equations of motion for a uniform 1-dimensional chain of particles of mass m with

nearest-neighbor interaction can be written as

$$m \frac{d^2 Q_n}{dt^2} = -\varphi'(Q_n - Q_{n-1}) + \varphi'(Q_{n+1} - Q_n) \\ (n = \dots, 1, 2, \dots).$$

It was shown that a lattice with a nonlinear interaction of the form

$$\varphi(r) = e^{-r} + r + \text{const.}$$

admits solutions in closed form. Later, it was shown that this lattice (the **exponential lattice** or the **Toda lattice**) is a completely **integrable system**.

It is convenient to introduce s_n , which is the generalized momentum canonically conjugate to the mutual displacement $r_n = Q_{n+1} - Q_n$. For a lattice with exponential interaction,

$$e^{-r_n} - 1 = ds_n/dt,$$

and if we introduce

$$S_n = \int^t s_n dt,$$

then the equations of motion can be written as

$$\log(1 + d^2 S_n / dt^2) = S_{n+1} + S_{n-1} - 2S_n,$$

and the displacements are given by

$$Q_n = S_n - S_{n+1}.$$

We have a solution

$$S_n = \log \{1 + e^{2(\alpha n + \beta t + \delta)}\},$$

where α and δ are arbitrary constants and $\beta = \pm \sinh \alpha$. The associated wave

$$e^{-r_n} - 1 = \beta^2 \text{sech}^2(\alpha n + \beta t + \delta)$$

represents a solitary wave or soliton.

The multisoliton (N -soliton) solution is given by

$$S_n = \log \det \Psi_n,$$

where Ψ_n is an $N \times N$ matrix whose elements are

$$(\Psi_n)_{jk} = \delta_{jk} + c_j c_k \frac{(Z_j Z_k)^{n+1}}{1 - Z_j Z_k} e^{(\beta_j + \beta_k)t}$$

with

$$Z_j = \pm e^{-\alpha_j},$$

$$\beta_j = \mp \sinh \alpha_j.$$

Asymptotically the wave reduces as $t \rightarrow \mp \infty$ to an assembly of solitons

$$e^{-r_n} - 1 = \sum_{j=1}^N \beta_j^2 \text{sech}^2(\alpha_j n + \beta_j t + \delta_j^\mp).$$

B. Conserved Quantities

The equations of motion for a periodic exponential lattice of N particles can be written

in a **Lax representation** $dL/dt = BL - LB$, where L and B are $N \times N$ matrices with the elements

$$L_{nn} = b_n, \quad L_{n,n+1} = L_{n+1,n} = a_n,$$

$$L_{1,N} = L_{N,1} = a_N,$$

$$B_{n,n+1} = -B_{n+1,n} = -a_n,$$

$$B_{1,N} = -B_{N,1} = a_N$$

(the other elements of L and B are all zero), with

$$a_n = \frac{1}{2} e^{-(Q_{n+1} - Q_n)/2},$$

$$b_n = \frac{1}{2} P_n \quad (P_n = dQ_n/dt).$$

The eigenvalues λ of L can be shown to be independent of time, and so the motion of the lattice is a spectrum-preserving deformation. Now, if we define $\{I_i\}$ by

$$\det(\lambda I - L) = \lambda^N + \lambda^{N-1} I_1 + \dots + \lambda I_{N-1} + I_N,$$

the $n\{I_i\}$ are polynomials of a_n and b_n . These are constants of motion that were discovered independently by M. Hénon and H. Flaschka. Thus the lattice has N conserved quantities; I_1 is related to the total momentum and I_2 to the total energy, but higher-index conserved quantities have no physical interpretation.

C. Method of Integration

Let L and B be the infinite matrices obtained from the foregoing ones in the limit $N \rightarrow \infty$. The eigenvalues λ of the equation

$$L\varphi = \lambda\varphi$$

are independent of time, and the time evolution of φ is given by the equation

$$d\varphi/dt = B\varphi.$$

If the motion in the lattice is restricted to a finite region, we can clearly speak of the scattering of the wave φ due to the deformation in the lattice. For a given initial motion $Q_n(0)$ and $P_n(0)$, or $L(0)$, we calculate the initial scattering data of asymptotic form $\varphi \sim z^n$ ($n \rightarrow \infty$). The **scattering data** consist of the reflection coefficient $R(z)$, the bound state eigenvalue $\lambda_j = -(z_j + z_j^{-1})/2$ ($\lambda_j < -1$ or $\lambda_j > 1$), and the coefficient c_j of the normalized bound state eigenfunction of asymptotic form $c_j z_j^n$ for $n \rightarrow \infty$.

From the initial data and the equations of motion for $n \rightarrow +\infty$, we get the scattering data at a later time t . In effect, we construct the kernel

$$F(m) = \frac{1}{2\pi i} \oint R(z, 0) e^{-(z - z^{-1})t} z^{m-1} dz \\ + \sum_j c_j^2(0) e^{-(z_j - z_j^{-1})t} z_j^m$$

of the discrete integral equation (**Gelfand-Levitan-Morchenko equation**)

$$\kappa(n, m) + F(n + m) + \sum_{n'=n+1}^{\infty} \kappa(n, n') F(n' + m) = 0, \\ m \geq n + 1.$$

After solving this equation for $\kappa(n, m)$, we calculate $K(n, n)$, given by

$$\frac{1}{[K(n, n)]^2} = 1 + F(2n) + \sum_{n'=n+1}^{\infty} \kappa(n, n') F(n' + n).$$

Then the initial value problem is solved in the form

$$e^{-(Q_n - Q_{n-1})} = \left[\frac{K(n, n)}{K(n-1, n-1)} \right]^2.$$

The solution can be given as $dQ_n/dt = s_n - s_{n+1}$, with

$$s_n = \kappa(n-1, n).$$

The simplest case $R(z) = 0$ yields the multi-soliton solution.

For the periodic case also, eigenvalues of the equations $L\varphi = \mu\varphi$ and $d\varphi/dt = B\varphi$, under suitable boundary conditions and for certain initial data, give sufficient information to construct a solution to the initial value problem. Such a method of obtaining a general solution for the periodic lattice was developed by E. Date and S. Tanaka, and independently by M. Kac and P. Van Moerbeke. Following Date and Tanaka, the solution can be written in terms of the multivariable theta function, or the Riemann theta function \mathfrak{J} , as

$$P_n = \frac{d}{dt} \log \frac{\mathfrak{J}(\alpha n + \beta t + \delta)}{\mathfrak{J}(\alpha(n+1) + \beta t + \delta)} + \text{const.},$$

where α , β , and δ are certain vector constants.

There has been much activity recently toward interpreting the integrability of the Toda lattice in terms of Lie algebras (B. Kostant [9]).

References

- [1] E. Fermi, J. Pasta, and S. Ulam, Studies of non-linear problems, Los Alamos Report LA-1940 (1955); Collected papers of Enrico Fermi, vol. II, Univ. of Chicago Press, 1965, 978–988.
- [2] M. Toda, Vibration of a chain with non-linear interaction, J. Phys. Soc. Japan, 22 (1967), 431–436.
- [3] M. Toda, Studies of a nonlinear lattice, Phys. Rep., 18C (1975), 1–123.
- [4] M. Hénon, Integrals of the Toda lattice, Phys. Rev., B9 (1974), 1921–1923.
- [5] H. Flaschka, On the Toda lattice. I, Existence of integrals, Phys. Rev., B9 (1974), 1924–1925; II, Inverse-scattering solution, Prog. Theor. Phys., 51 (1974), 703–716.

[6] E. Date and S. Tanaka, Analogue of inverse scattering theory for the discrete Hill's equation and exact solutions for the periodic Toda lattice, Prog. Theor. Phys., 55 (1976), 457–465.

[7] M. Kac and P. Van Moerbeke, A complete solution of the periodic Toda problem, Proc. Nat. Acad. Sci. US, 72 (1975), 2879–2880.

[8] M. Toda, Theory of nonlinear lattices, Springer, 1981. (Original in Japanese, 1978.)

[9] B. Kostant, The solution to a generalized Toda lattice and representation theory, Advances in Math., 34 (1979), 195–338.

288 (XIII.10) Nonlinear Ordinary Differential Equations (Global Theory)

A. General Remarks

Many well-known functions (with the notable exception of the Γ -function), such as the exponential, trigonometric, elliptic, and automorphic functions, satisfy ordinary differential equations of simple forms. For the purpose of finding new transcendental functions, P. Painlevé initiated the systematic study of the equation

$$F(x, y, y', \dots, y^{(n)}) = 0 \quad (1)$$

in the complex domain. To investigate the solution in its whole domain of definition, he assumed that F is a polynomial in $y, y', \dots, y^{(n)}$ whose coefficients are analytic in x . Such an equation is called an **algebraic differential equation**. If F is linear in $y^{(n)}$, then (1) is written as

$$y^{(n)} = \frac{P(x, y, y', \dots, y^{(n-1)})}{Q(x, y, y', \dots, y^{(n-1)})}, \quad (2)$$

where P and Q are polynomials in $y, y', \dots, y^{(n-1)}$ with coefficients that are analytic functions of x . Equation (2) is called a **rational differential equation**.

If F is linear in $y, y', \dots, y^{(n)}$, i.e., (1) is a linear differential equation, and if the coefficient of $y^{(n)}$ is 1, then singular points of solutions are situated at the singular points of the coefficients (\rightarrow 254 Linear Ordinary Differential Equations (Local Theory)). If F is not linear, then singular points of solutions of (1) are divided into two categories, one consisting of those points whose positions are determined by the equation itself and are independent of individual solutions, and the other consisting of those points whose positions depend on the

choice of particular solutions. In other words, the singularities of the first category appear independently of the choice of arbitrary constants involved in the general solution, while those of the second category depend on the choice of the arbitrary constants. The former are called **fixed singularities** and the latter **movable singularities**. The linear differential equation (1) has fixed singularities only, which are situated at the singularities of the coefficients. In the same way, [†]branch points of solutions can be classified into two kinds, **fixed branch points** and **movable branch points**.

B. Algebraic Differential Equations of the First Order

Consider the equation

$$y' = \frac{P(x, y)}{Q(x, y)}, \quad (3)$$

where P and Q are relatively prime polynomials in x, y . The fixed singular points of (3) are defined to be points ξ, ξ' with the following properties: (i) $Q(\xi, y) \equiv 0$. (ii) $Q(\xi', y) \not\equiv 0$, $P(\xi', y) = Q(\xi', y) = 0$ have a root $y = \eta'$. (iii) If, in (ii), we substitute $1/z$ for y and if the same relation (ii) holds for $x = \xi$ and $z = 0$, we count such a value ξ_1 as a fixed singular point. (iv) We transform equation (3) by setting $x = 1/t$. If the value $t = 0$ satisfies (i) or (ii) for this transformed equation, we count ∞ as a singular point ξ or ξ' accordingly. The points ξ, ξ' are, in general, [†]transcendental singularities of solutions, and the points ξ' cannot be [†]essential singularities of solutions but may be [†]ordinary transcendental singularities. A singular point of a solution different from ξ and ξ' is an [†]algebraic singular point, and for any point distinct from ξ and ξ' , equation (3) admits a solution with an algebraic singularity at this point. A necessary and sufficient condition that (3) has no movable branch point is that (3) be a [†]Riccati equation.

Consider the algebraic differential equation of the first order

$$F(x, y, y') = 0. \quad (4)$$

After defining the fixed singular points ξ and ξ' of (4), where the algebraic function of x, y defined by (4) has bad singularities, Painlevé proved that movable singularities of solutions are algebraic. Before Painlevé's work, L. Fuchs gave a necessary and sufficient condition that (4) has no movable branch points, and then H. Poincaré showed that if this condition is satisfied, then (4) is either reducible to the Riccati equation if $g = 0$, integrable by the use of elliptic functions if $g = 1$, or algebraically integrable if $g > 1$, where, with x the independent vari-

able, g denotes the [†]genus of the algebraic curve defined by (4). Painlevé found that there were gaps in the proofs of Fuchs and Poincaré and completed these by proving his theorem and the following one: Let $\varphi(x, y_0, x_0)$ be the solution of (4) satisfying the initial condition $y(x_0) = y_0$. Let \bar{x}, \bar{x}_0 be points different from ξ and ξ' , and let L be a curve connecting \bar{x}_0 to \bar{x} and not passing through any ξ or ξ' . If we denote by $\varphi_L(\bar{x}, y_0, \bar{x}_0)$ the value at $x = \bar{x}$ of the branch obtained by continuing $\varphi(x, y_0, \bar{x}_0)$ analytically in a neighborhood of L and regard $\varphi_L(\bar{x}, y_0, \bar{x}_0)$ as a function of y_0 , then $\varphi_L(\bar{x}, y_0, \bar{x}_0)$ coincides, in a neighborhood of every point $y_0 = b$, with several branches of an [†]algebraic function of y_0 . Painlevé studied the case when the general solution is finitely many-valued and gave a condition for $\varphi(x, y_0, x_0)$ to be an algebraic function of y_0 .

When equation (4) does not contain x explicitly, no movable branch points appear if and only if all solutions are single-valued, and then the solutions are expressible in terms of rational, exponential, and elliptic functions. Such an equation is called a **Briot-Bouquet differential equation**.

J. Malmquist proved, by using P. Boutroux's method of studying the behavior of solutions in the neighborhood of a fixed singularity, that if equation (4) admits at least one solution that has an essential singularity and is finitely many-valued and free from movable branch points around this singularity, then (4) is an equation without movable branch points. If (4) admits a solution that is a finitely many-valued transcendental function, then an algebraic transformation may be applied to (4) so that it will become an equation without movable branch points. It is an immediate consequence of the first assertion that if (3) admits a solution that has an essential singularity and is finitely many-valued and free from movable branch points around the singularity, then (3) is a Riccati equation.

Later, equations (3) and (4) were studied by M. Hukuhara, K. Yosida, T. Satō, T. Kimura, and T. Matuda. The following results are due to Kimura. If a solution $\varphi(x)$ of (3) has an essential singularity at $x = \xi$, then, in an arbitrary neighborhood of ξ , $\varphi(x)$ assumes every value with the exception of the roots of $P(\xi, y) = 0$. If (3) is not a Riccati equation, it is determined by a finite number of algebraic processes whether or not (3) admits a solution that has an essential singularity at $x = \xi$ and has no movable branch point around ξ . If (3) admits such a solution, then the singularity is a [†]logarithmic branch point. For an essential singularity of a solution there exists, in general, a direction similar to a [†]Julia's direction, which was investigated by Hukuhara and

Kimura. Matuda studied in detail the behavior of solutions as x tends to ξ along a half-line and concluded that, except for some special cases, any solution tends to a certain value as x tends to ξ along a half-line. To obtain algebraic solutions C. Briot and J. C. Bouquet devised a method similar to [†]Puiseux expansion in the theory of algebraic functions. Hukuhara improved their method and succeeded in reducing (3) to several differential equations of standard forms in a neighborhood of $x = \xi$. This enables us to apply the local theory to the global study. Hukuhara's method was used by Kimura and Matuda to obtain the results discussed in this paragraph.

C. Algebraic Differential Equations of the Second Order

For second-order algebraic differential equations we can pose the same problems as for first-order equations: When do these equations have single-valued or finitely many-valued general solutions? What new transcendental functions are needed to integrate such equations? These problems, studied by E. Picard and Painlevé, are difficult because of the existence of movable transcendental singularities. However, Painlevé succeeded in determining rational differential equations of the second order without movable branch points. Such equations, with the exception of those that are integrated by the use of solutions of the first-order and linear differential equations, can be transformed by rational transformations into one of the following six differential equations:

$$(I) \quad y'' = 6y^2 + x,$$

$$(II) \quad y'' = 2y^3 + xy + \alpha,$$

$$(III) \quad y'' = \frac{y'^2}{y} - \frac{y'}{x} + \frac{\alpha y^2 + \beta}{x} + \gamma y^3 + \frac{\delta}{y},$$

$$(IV) \quad y'' = \frac{y'^2}{2y} + \frac{3y^3}{2} + 4xy^2 + 2(x^2 - \alpha)y + \frac{\beta}{y},$$

$$(V) \quad y'' = y'^2 \left(\frac{1}{2y} + \frac{1}{y-1} \right) - \frac{y'}{x} + \frac{(y-1)^2}{x^2} \left(\alpha y + \frac{\beta}{y} \right) + \frac{\gamma y}{x} + \frac{\delta y(y+1)}{y-1},$$

$$(VI) \quad y'' = \frac{y'^2}{2} \left(\frac{1}{y} + \frac{1}{y-1} + \frac{1}{y-x} \right) - y' \left(\frac{1}{x} + \frac{1}{x-1} + \frac{1}{y-x} \right) + \frac{y(y-1)(y-x)}{x^2(x-1)^2} \\ \times \left\{ \alpha + \frac{\beta x}{y^2} + \frac{\gamma(x-1)}{(y-1)^2} + \frac{\delta x(x-1)}{(y-x)^2} \right\},$$

where α , β , γ , and δ are constants. These equa-

tions and their solutions are called **Painlevé equations** and **Painlevé transcendental functions**, respectively. Equation (VI) was discovered by B. Gambier, who found an omission in Painlevé's calculations. All solutions of (I) are single-valued, and their properties were investigated by Boutroux. The solutions of (VI) have, in general, logarithmic branch points at $x=0, 1, \infty$ and were studied by R. Garnier.

The case when the equation is of degree 2 with respect to y'' was studied by Malmquist and F. Tricomi.

The following facts are known concerning movable transcendental singularities of the rational equation $y'' = P(x, y, y')/Q(x, y, y')$, where P and Q are relatively prime polynomials (Kimura). Let p, q be the degrees of P and Q with respect to y' . If $p > q + 2$, then any solution $\varphi(x)$ admits no movable essential singularities, but its derivative $\varphi'(x)$ may admit such singularities. If $p > q + 2$ and Q is not decomposable as $Q_1(x, y)Q_2(x, y, y')$, then neither φ nor φ' admits movable essential singularities. If $p \leq q + 2$, then both φ and φ' may have movable essential singularities. If $p \leq q + 2$ and Q is not decomposable as above, then φ' has no movable essential singularities. If, for a solution $\varphi(x)$, $x=a$ is a movable essential singularity of $\varphi(x)$ or $\varphi'(x)$, the $\varphi(x)$ or $\varphi'(x)$ assumes all values other than a finite number of exceptional values in an arbitrary neighborhood of $x=a$. If $p > q + 2$, then every solution possesses [†]Iversen's property, and hence the set of movable singularities is not a continuum.

D. Higher-Order Equations and Other Equations

Painlevé's method of obtaining the second-order equations without movable branch points is applicable to higher-order equations. The determination of third-order equations without movable branch points was attempted by Painlevé, J. Chazy, and Garnier by the use of this method, but is not yet complete. Chazy studied in detail an equation of the form

$$y''' = \frac{(1-1/n)y''^2}{y'} + b(y)y'y'' + c(y)y'^3$$

and showed that when $n = -2$ and $b(y) = 0$, [†]Fuchsian and [†]Kleinian functions are obtained as solutions (\rightarrow 32 Automorphic Functions).

R. Fuchs, a son of L. Fuchs, derived equation (VI), at almost the same time as Gambier, from the study of [†]monodromy groups. He showed that the monodromy group of the

equation

$$\frac{1}{z} \frac{d^2 z}{dt^2} = \left(\frac{\alpha}{t^2} + \frac{\beta}{(t-1)^2} + \frac{\gamma}{(t-x)^2} + \frac{\delta}{t(t-1)} + \frac{3}{4(t-y)^2} + \frac{a}{t(t-1)(t-x)} + \frac{b}{t(t-1)(t-y)} \right)$$

remains invariant as the singularity x varies if and only if α , β , γ , and δ remain constant; the singularity y , considered as a function of x , satisfies equation (VI); and a and b are rational functions of x and y and y' . Investigating a second-order linear differential equation of Fuchsian type with \dagger regular singularities $0, 1, \infty, x_1, \dots, x_n, y_1, \dots, y_n$ where y_1, \dots, y_n are apparent fixed singularities, Garnier was led, under the hypothesis that the monodromy group of this equation remains invariant as x_1, \dots, x_n vary, to a \dagger completely integrable system of partial differential equations and showed that a symmetric function of y_1, \dots, y_n , considered as a function of any one of x_1, \dots, x_n , satisfies an equation without movable branch points (\rightarrow 253 Linear Ordinary Differential Equations (Global Theory)).

For nonalgebraic equations, movable branch points appear even in the first-order case. Kimura obtained a sufficient condition that for an equation $F(x, y, y') = 0$, where F is a polynomial of y' with meromorphic coefficients in x, y , every solution has Iversen's property in a domain of the complex plane.

It was shown by O. Hölder that the Γ -function satisfies no algebraic differential equation. Also, if a function meromorphic in the unit circle is of \dagger order ∞ , then it satisfies no algebraic differential equation.

References

- [1] P. Painlevé, *Oeuvres de Paul Painlevé*, 1, 2, 3, Centre Nat. Recherche Sci. France, 1972, 1974, 1975.
- [2] P. Boutroux, *Leçons sur les fonctions définies par les équations différentielles du premier ordre*, Gauthier-Villars, 1908.
- [3] E. L. Ince, *Ordinary differential equations*, Longmans-Green, 1927.
- [4] L. Bieberbach, *Theorie der gewöhnlichen Differentialgleichungen*, Springer, second edition, 1965.
- [5] M. Hukuhara, T. Kimura, and T. Matuda, *Equations différentielles ordinaires du premier ordre dans le champ complexe*, Publ. Math. Soc. Japan, 1961.
- [6] J. Malmquist, *Sur les fonctions à un nombre fini de branches définies par les équations*

différentielles du premier ordre, Acta Math., 36 (1913), 297–343; 74 (1941), 175–196.
 [7] E. Hille, *Ordinary differential equations in the complex domain*, Wiley, 1976.

289 (XIII.9) Nonlinear Ordinary Differential Equations (Local Theory)

A. General Remarks

Consider a system of n differential equations

$$dy_j/dx = f_j(x, y_1, \dots, y_n), \quad j = 1, \dots, n, \quad (1)$$

where the f_j are analytic functions of x, y_1, \dots, y_n . To simplify the notation, we use vector notation \mathbf{y} instead of (y_1, \dots, y_n) . If all the f_j are holomorphic at a point $(x, \mathbf{y}) = (a, \mathbf{b})$, there exists one and only one solution $\mathbf{y}(x)$ for (1) such that $\mathbf{y} \rightarrow \mathbf{b}$ as $x \rightarrow a$ and $\mathbf{y}(x)$ is holomorphic at $x = a$. We say that a point (a, \mathbf{b}) is a **singular point** of the system (1) if it is a singular point of f_j for some j . The well-known \dagger Cauchy's existence theorem can no longer be applied to the case when (a, \mathbf{b}) is a singular point of system (1). In this case, the following three problems arise naturally: (i) to determine whether solutions $\mathbf{y}(x)$ such that $\mathbf{y}(x) \rightarrow \mathbf{b}$ as $x \rightarrow a$ exist, and if they exist, to determine the number of independent solutions; (ii) to construct analytic expressions for solutions $\mathbf{y}(x)$ such that $\mathbf{y}(x) \rightarrow \mathbf{b}$ as $x \rightarrow a$ or, in a slightly more general way, analytic expressions for bounded solutions $\mathbf{y}(x)$ such that the values of $(x, \mathbf{y}(x))$ stay in a neighborhood of the singular point (a, \mathbf{b}) ; (iii) to investigate the properties of these solutions. These three problems are called **local problems**, since only those solutions in a neighborhood of the singular point (a, \mathbf{b}) are considered. However, even when $n = 1$, the study of local problems is very difficult except for the case of singular points of particular types at which the functions f_j are meromorphic.

When $n > 1$, the problem becomes even harder; research on this case lags that for $n = 1$. In the subsequent discussion we assume without loss of generality that $a = 0, \mathbf{b} = \mathbf{0}$.

B. The Case of a Single Equation

Consider the equation

$$dy/dx = Y(x, y)/X(x, y), \quad (2)$$

where X and Y are holomorphic functions of (x, y) at $(0, 0)$, $x, y \in \mathbb{C}$. When $X(0, 0) = 0$ and

Nonlinear ODEs (Local Theory)

$Y(0,0) \neq 0$, we can rewrite equation (2) in the form $dx/dy = X(x,y)/Y(x,y)$, and we see that (i) if $X(0,y) \neq 0$, equation (2) has one and only one solution $y(x)$ which is algebraic at $x=0$ and tends to 0 as $x \rightarrow 0$; (ii) if $X(0,y) \equiv 0$, there is no solution of equation (2) such that $y \rightarrow 0$ as $x \rightarrow 0$.

The case where $X(0,0)=0$ and $Y(0,0)=0$ was first studied by C. A. A. Briot and J. C. Bouquet. In order to obtain algebraic solutions they introduced a method similar to [†]Puiseux expansion in the theory of algebraic functions. A. R. Forsyth and J. Malmquist studied a problem of reduction for equation (2) by using the Briot-Bouquet method. The theory of reduction was completed by M. Hukuhara, who divided a neighborhood of $(0,0)$ into a finite number of subdomains in such a way that (i) the union of these subdomains covers the given neighborhood of $(0,0)$ completely; (ii) in each of these subdomains equation (2) takes one of eight canonical forms. Hukuhara investigated the properties of the solutions for these canonical forms. Among them, the following two are well known as the **Briot-Bouquet differential equations**:

$$x dy/dx = f(x, y), \quad f(0,0)=0, \quad (3)$$

$$x^{\sigma+1} dy/dx = f(x, y),$$

$$f(0,0)=0, \quad \sigma \geq 1 \text{ is an integer.} \quad (4)$$

C. Systems of Differential Equations

When y is a vector with components (y_j) , equations (3) and (4) can be written as

$$x dy_j/dx = f_j(x, y_1, \dots, y_n), \quad j=1, \dots, n, \quad (5)$$

$$x^{\sigma+1} dy_j/dx = f_j(x, y_1, \dots, y_n), \quad j=1, \dots, n. \quad (6)$$

After Briot and Bouquet, the singular points of these types were studied by many authors, including Poincaré, E. Picard, H. Dulac, Malmquist, W. J. Trjitzinski, and Hukuhara. A method of constructing solutions of these equations consists of two parts: (i) formally transforming the given equations into simplified or reduced equations of the simplest possible form by applying a formal transformation of the type

$$y = \sum p_{kl} x^k z^l \quad (7)$$

or

$$y_j = \sum p_{k_0 k_1 \dots k_n} x^{k_0} z_1^{k_1} \dots z_n^{k_n}, \quad j=1, 2, \dots, n; \quad (8)$$

(ii) verifying convergence or the validity of [†]asymptotic expansions for the **formal solutions** of the given equations, which are obtained by substituting bounded solutions of the equations satisfied by z or z_j into (7) or (8).

By studying these analytic expressions, the properties of the solutions can be clarified.

D. Properties of Solutions of Briot-Bouquet Differential Equations

For equation (3), the character of the simplified equation depends on the value of $\lambda = f_y(0,0)$. We have the following four cases: (i) λ is neither 0 nor negative. A suitable formal transformation (7) changes (3) to

$$x dz/dx = \lambda z + bx^\lambda.$$

In particular, if λ is not equal to a positive integer, then $b=0$. The double power series (7) is uniformly convergent. There exists a function $\varphi(x, z)$ of (x, z) holomorphic at $(0,0)$ such that $y = \varphi(x, x^\lambda(b \log x + C))$ is a general solution of (3) with an integration constant C . (ii) $\lambda=0$. The equation satisfied by z takes the form

$$x dz/dx = z^{m+1}(b + b'z^m), \quad m \geq 1.$$

If $b=0$, b' necessarily vanishes and $x=0$ is a holomorphic point of (3). A general solution is given by

$$Z(x) = (C - mb \log x)^{-1/m}, \quad b \neq 0, \quad b' = 0,$$

and

$$Z(x) = \left(\frac{b'}{b} \alpha \left(\frac{mb^2}{b'} \log \frac{1}{x} + C \right) \right)^{-1/m}, \quad bb' \neq 0.$$

Here, $\zeta = \alpha(t)$ is the branch (of the inverse function of $\zeta - \log \zeta = t$) such that $\alpha(t) - t \log t - t \rightarrow 0$ as $t \rightarrow \infty$. Then there exists a holomorphic function $\varphi(x, z)$ of (x, z) for $|x| < \delta$, $|m \arg z + \arg b| < 3\pi/2 - \varepsilon$, $|z| < \Delta$ such that $\varphi(x, Z(x))$ is a general solution of (3). (iii) λ is a negative rational number $-\mu/\nu$. The equation satisfied by z is written as

$$\nu x dz/dx = z(-\mu + b(x^\mu z^\nu)^\mu + b'(x^\mu z^\nu)^{2\mu}).$$

A general solution has the form

$$Z(x) = x^\lambda (C - mb \log x)^{-1/m\nu}, \quad b' = 0,$$

and

$$Z(x) = x^\lambda \left(\frac{b'}{b} \alpha \left(\frac{\nu mb^2}{b'} \log \frac{1}{x} + C \right) \right)^{-1/m\nu},$$

$$bb' \neq 0.$$

In this case, there exists a holomorphic function $\varphi(x, z)$ of (x, z) for $|m\mu \arg x + m\nu \arg z + \arg b \mp \pi/2| < \pi - \varepsilon$, $|x| < \delta$, $|z| < \Delta$ such that $\varphi(x, Z(x))$ is a general solution of (3). M. Iwano expressed this solution in the form $\psi((x^\mu Z(x))^\mu, x, Z(x))$, where $\psi(w, x, z)$ is holomorphic for $|\arg w \mp \frac{\pi}{2}| < \pi - \varepsilon$, $0 < |w| < \delta$, $|x| < \Delta$, $|z| < \Delta$. In particular, if $b=b'=0$, $\varphi(x, z)$ is holomorphic at $(0,0)$. (iv) λ is a negative

irrational number. The equation in z has the form

$$x \, dz/dx = \lambda z.$$

The formal transformation (7) may either diverge or converge. Dulac proved that if (7) diverges and if there exists a solution $y(x)$ such that $y(x) \rightarrow 0$ as $x \rightarrow 0$ along a suitable path L , then $|x^\alpha y(x)^\beta \arg x| \rightarrow \infty$ and $|x^\alpha y(x)^\beta \arg y(x)| \rightarrow \infty$ as $x \rightarrow 0$, $x \in L$ for any α and β . However, the existence of such a solution is not yet verified. C. L. Siegel proved that if λ satisfies certain inequalities, the formal transformation (7) is divergent. An example such that (7) is divergent was first given by Dulac. A very simple example for such a case was given by Y. Sibuya.

In the case $f(0, 0) = 0$ and $\lambda = f_y(0, 0) \neq 0$ for (4), the most complete result was obtained by Hukuhara, namely: (i) A suitable formal transformation (7) reduces equation (4) to

$$x^{\sigma+1} \, dz/dx = z(\alpha_0 + \alpha_1 x + \dots + \alpha_n x^n), \quad \alpha_0 = \lambda.$$

A general solution is given by $Z(x) = C \cdot x^{\alpha_0} e^{\Lambda(x)}$, where $\Lambda(x)$ is a polynomial in $1/x$ of degree σ . (ii) A general solution of (4) is expressed by a uniformly convergent power series of the form $\sum \varphi_k(x) Z(x)^k$, where the $\varphi_k(x)$ are holomorphic functions of x for a certain sectorial neighborhood of $x=0$ and have asymptotic expansions $\sum p_{jk} x^k$ as $x \rightarrow 0$.

Assume $f_j(0, 0, \dots, 0) = 0$ for equations (5). Put $\lambda_{jk} = \partial f_j / \partial y_k(0, 0, \dots, 0)$. Denote by $\lambda_1, \dots, \lambda_n$ the eigenvalues of an $n \times n$ matrix with elements $\{\lambda_{jk}\}$. Then (i) if an angle ω can be chosen so that for some $m \leq n$, all of $|\omega|$, $|\arg \lambda_1 - \omega|, \dots, |\arg \lambda_m - \omega|$ are less than $\pi/2$, equations (5) possess solutions that are expressed as uniformly convergent $(m+1)$ -tuple power series of x , $Z_1(x), \dots, Z_m(x)$:

$$\sum p_{jk_0 k_1 \dots k_m} x^{k_0} Z_1^{k_1} \dots Z_m^{k_m}.$$

Here the $Z_k(x)$ are general solutions of the simplified equations and have the expression

$$Z_k(x) = x^{\lambda_k} (C_k + \text{a polynomial of } C_1, \dots, C_{k-1}, \log x), \quad k = 1, 2, \dots, m.$$

(ii) Moreover, Iwano extended the result of (i) as follows: When there exists one and only one zero among the other $n-m$ eigenvalues, equations (5) have solutions, depending on $m+1$ arbitrary constants, that are expressed as $(m+1)$ -tuple uniformly convergent power series

$$\sum p_{jk_0 k_1 \dots k_m} (Z_{m+1}(x)) x^{k_0} Z_1^{k_1} \dots Z_m^{k_m}.$$

Here the coefficients $p_{jk_0 k_1 \dots k_m}(z_{m+1})$ are holomorphic functions of z_{m+1} in a sectorial neighborhood of $z_{m+1} = 0$ and admit asymptotic expansions in powers of z_{m+1} as $z_{m+1} \rightarrow 0$. In this case, the functions $Z_1(x), \dots, Z_{m+1}(x)$ can-

not generally be integrated by quadratures, but as $Z_{m+1}(x) \rightarrow 0$, the power series expansions coincide with the expressions obtained in (i). (iii) In the case when the Jacobian matrix (λ_{jk}) is the zero matrix, Iwano constructed, under additional assumptions, a convergent analytic expression of a general solution for equations (5).

Let the right-hand side of (6) be holomorphic at $(0, 0, \dots, 0)$. In 1939 Trjitzinski proved the existence of solutions that admit asymptotic expansions in powers of n arbitrary constants. In 1940 and 1941, Malmquist proved, under strong conditions, the existence of solutions that are expressed as uniformly convergent power series of $Z_\alpha(x), \dots, Z_\beta(x)$: $\sum p_{jk_\alpha \dots k_\beta}(x) Z_\alpha(x)^{k_\alpha} \dots Z_\beta(x)^{k_\beta}$. Here the $Z_k(x)$ are polynomials of x and $\log x$ of the form

$$Z_k(x) = e^{\Lambda_k(x)} x^{\lambda_k} (C_k + \text{a polynomial of } C_{k-1}, \dots, C_\alpha, \log x), \quad k = \alpha, \dots, \beta,$$

where the coefficients admit asymptotic expansions in powers of x . Trjitzinski's result is contained in Malmquist's result as a special case. On the other hand, under much weaker assumptions than Malmquist's, Hukuhara solved a problem on the formal simplification of equations (6) and formal solutions. Iwano improved Hukuhara's result on formal solutions and discussed the convergence of formal solutions under weaker conditions than Malmquist's (*Ann. Mat. Pura Appl.*, 1957, 1959). The $p_{jk_\alpha \dots k_\beta}(x)$ are holomorphic functions of x in a sectorial neighborhood of $x=0$ and admit asymptotic expansions in powers of x as $x \rightarrow 0$. The angle of the sector in which the asymptotic expansions are valid is largest for Iwano's method.

If $\sum \sigma_j < n$, equations of the form

$$x^{\sigma_j} dy_j/dx = f_j(x, y_1, \dots, y_n), \quad j = 1, \dots, n,$$

possess at least $n - \sum \sigma_j$ solutions that are holomorphic for $|x| < \delta$ (R. W. Bass, *Amer. J. Math.*, 77 (1955)). This result is analogous to those obtained by O. Perron, F. Lettenmyer, and Hukuhara and Iwano in the linear case (\rightarrow 254 Linear Ordinary Differential Equations (Local Theory)).

E. Singular Perturbations

The terms on the right-hand side of the nonlinear differential equations

$$e^{\sigma_j} dy_j/dx = f_j(x, y_1, \dots, y_n, \varepsilon), \quad j = 1, 2, \dots, n, \quad (9)$$

are holomorphic functions of (x, y, ε) for $|x| < a$, $\|y\| < b$, $0 < |\varepsilon| < c$, $|\arg \varepsilon| < d$, and admit uniformly convergent expansions in powers of y with coefficients asymptotically developable in

powers of ε . The σ_j are nonnegative integers. W. R. Wasow, W. A. Harris, Sibuya, and Iwano and T. Saito discussed problems on constructing asymptotic or convergent expansions for bounded solutions that are dependent on several arbitrary constants.

In equations (9), the f_j and $\partial f_j / \partial y_k$ are continuous functions of (x, y, ε) for $-\infty < x < +\infty$, $\|y\| < b$, $|\varepsilon| < c$, and periodic functions of period T with respect to x . Moreover, assume that a system of degenerate algebraic equations

$$0 = f_j(x, y_1, \dots, y_n, 0), \quad j = 1, \dots, n, \quad (10)$$

has a periodic solution $y_j = p_j(x)$ of period T for $-\infty < x < +\infty$. Then if equations (9) have periodic solutions $y_j = p_j(x, \varepsilon)$ of period T such that $y_j \rightarrow p_j(x)$ as $\varepsilon \rightarrow 0$, the $p_j(x, \varepsilon)$ are called **singular perturbations** of $p_j(x)$ for equations (9). Concerning this problem, see I. M. Volk (*Prikl. Mat. Mekh. SSSR*, 10 (1946)). The work of Wasow (1950) on a single equation

$$\varepsilon^\sigma y^{(n)} = f(x, y, y', \dots, y^{(m)}, \varepsilon), \quad \sigma > 0, \quad n > m \geq 0, \quad (11)$$

is remarkable.

References

- [1] H. Dulac, Points singuliers des équations différentielles, *Mémoires. Sci. Math.*, 61, Gauthier-Villars, 1934.
- [2] E. Picard, *Traité d'analyse III*, Gauthier-Villars, 1896.
- [3] E. L. Ince, *Ordinary differential equations*, Dover reprint, 1956.
- [4] M. Hukuhara, T. Kimura, and T. Matuda, *Equations différentielles ordinaires du premier ordre dans le champ complexe*, Math. Soc. Japan, 1961.
- [5] E. Hille, *Ordinary differential equations in the complex domain*, Wiley, 1976.

290 (XIII.11) Nonlinear Oscillation

A. General Remarks

By **nonlinear oscillation** we usually mean oscillation described by periodic or \dagger almost periodic solutions of nonlinear ordinary differential equations. The theory of nonlinear oscillation is sometimes called **nonlinear mechanics**. In connection with oscillations in \dagger dynamical systems and electrical circuits, the theory of nonlinear oscillation has been studied intensively in the Soviet Union under the direction

of N. M. Krylov and N. N. Bogolyubov since around 1930, and after World War II research in this field became active in Western countries also.

Written as first-order systems, the differential equations in this theory take one of the forms

$$dx/dt = X(x) \quad (1)$$

or

$$dx/dt = X(x, t), \quad (2)$$

where x is a vector and t is a scalar. A differential equation of the form (1) is said to be **autonomous**. In a differential equation of the form (2), $X(x, t)$ is usually assumed to be periodic or almost periodic in t . In the former case, the differential equation (2) is said to be **periodic**, and in the latter case, **almost periodic**. Oscillations in physical systems are described mostly by periodic or almost periodic differential equations; therefore it is a principal problem in the theory of nonlinear oscillation to find a periodic or almost periodic solution of these differential equations. However, an oscillation described by a solution of a differential equation can be actually realized only when the solution is \dagger stable (\rightarrow 394 Stability) under a small variation of the initial value. Therefore it is important to investigate the stability of periodic or almost periodic solutions. In view of the fact that an actual phenomenon may be only approximately described by mathematical equations, sometimes it is necessary to require a certain stability of the system so that solutions of a periodic or almost periodic equation stay stable under a small variation of the equation itself. Such stability is called **structural stability**, the investigation of which is also important.

It may happen that a differential equation possesses neither a periodic solution nor an almost periodic solution, but that it has an almost periodic \dagger integral manifold (i.e., a manifold $x = f(t, \theta)$ in tx -space, where θ is a parameter, such that $f(t, \theta)$ is periodic or almost periodic in t and periodic in θ) containing the \dagger trajectory of the differential equation passing through an arbitrary point of the manifold. In this case, we can consider that a solution corresponding to a trajectory lying on the manifold describes an oscillation. Therefore it is important to find a periodic or almost periodic integral manifold of a differential equation and to investigate the stability of such a manifold.

The methods used most frequently in research on nonlinear oscillation are: (i) geometric methods, (ii) analytic methods, and (iii) numerical methods.

B. Linear Oscillations

Let S_1 and S_1^* be the spaces of all ω -periodic solutions of the ω -periodic linear system

$$dx/dt = A(t)x, \quad A(t + \omega) = A(t), \quad (3)$$

and its adjoint system, respectively. Then the space P_ω of continuous ω -periodic functions $\mathbf{R} \rightarrow \mathbf{R}^n$ has direct sum decompositions $P_\omega = S_1 + S_2$ and $P_\omega = S_1^* + S_2^*$ so that $(\mathbf{x}, \mathbf{y}) = 0$ for every $\mathbf{x} \in S_1$, $\mathbf{y} \in S_2$ or $\mathbf{x} \in S_1^*$, $\mathbf{y} \in S_2^*$, where $(\mathbf{x}, \mathbf{y}) = (1/\omega) \int_0^\omega \mathbf{x}(s)\mathbf{y}(s)ds$. Let $\{\xi^1, \dots, \xi^m\}$ and $\{\eta^1, \dots, \eta^m\}$ (m may be 0) be bases of S_1 and S_1^* orthonormal with respect to (\cdot, \cdot) . Then for $\mathbf{p} \in P_\omega$ there is a unique ω -periodic solution \mathbf{x} of

$$dx/dt = A(t)x + \mathbf{p}(t) \quad (4)$$

belonging to S_2 under the condition $(\eta^k, \mathbf{p}) = 0$ ($k = 1, \dots, m$), which is represented in the form $\mathbf{x} = G[\mathbf{p}]$ by a bounded linear operator $G: P_\omega \rightarrow P_\omega$. Also, if (3) and hence its adjoint system have no nontrivial ω -periodic solution, then (4) always has a unique ω -periodic solution given by $\mathbf{x} = G[\mathbf{p}]$, and this is true even if the ω -periodicity is replaced by almost periodicity. The almost periodic system (3) is said to be **regular** if (4) has almost periodic solution for an arbitrary almost periodic function $\mathbf{p}(t)$. A necessary and sufficient condition for (3) to be regular is that (3) induce an **exponential dichotomy**, that is, the solution space S of (3) have a direct sum decomposition $S = S_- + S_+$ such that $\|\mathbf{x}(t)\| \leq M e^{-\gamma|t-s|} \|\mathbf{x}(s)\|$ holds for $-\infty < s \leq t < \infty$ if $\mathbf{x} \in S_+$ and for $-\infty < t \leq s < \infty$ if $\mathbf{x} \in S_-$, where M and γ are positive constants. An autonomous (resp. periodic) system (3) is regular if and only if no \dagger characteristic roots (resp. characteristic exponents) have zero real parts.

C. Geometric Methods

Geometric methods are used frequently for finding a periodic solution in an autonomous or periodic case. In the autonomous case (1), a periodic solution describes a closed orbit (t is a parameter of a curve) in \mathbf{x} -space, which is usually called a **phase space**. The geometric method is used to show the existence and the stability of a closed orbit by investigating geometrically the behavior of orbits in the phase space. In such an approach, the properties of \dagger critical points and \dagger limit sets (\rightarrow 126 Dynamical Systems) are utilized frequently. This method is effective especially for 2-dimensional cases on the basis of the \dagger Poincaré-Bendixson theorem, and various results are given for a generalized **Liénard's** (or **Duffing's**)

differential equation $\ddot{x} + f(x)\dot{x} + g(x) = 0$, which includes the **van der Pol differential equation** $\ddot{x} - \lambda(1 - x^2)\dot{x} + x = 0$ [1, 2].

In the periodic case (2), let $\omega (> 0)$ be a period of $\mathbf{X}(\mathbf{x}, t)$ with respect to t and $\mathbf{x} = \varphi(t, \alpha)$ be a solution of (2) such that $\varphi(0, \alpha) = \alpha$. Then a periodic solution of (2) is given by $\mathbf{x} = \varphi(t, \alpha_0)$, where α_0 satisfies $\varphi(\omega, \alpha_0) = \alpha_0$. Thus the existence of a periodic solution can be shown by investigating geometrically the existence of a \dagger fixed point of the mapping $\mathbf{x} \rightarrow \mathbf{x}' = \varphi(\omega, \mathbf{x})$ in the phase space. In such an approach, \dagger Brouwer's fixed-point theorem is utilized frequently. In the mapping $\mathbf{x} \rightarrow \mathbf{x}'$, it may happen that there is no fixed point but that there exists an \dagger invariant manifold. In this case, we get a periodic integral manifold.

Geometric methods give information on qualitative properties, but usually not on quantitative properties like the shape of an oscillation. Thus these methods are in general not sufficient for the analysis of the phenomena met in practice.

D. Analytic Methods

At present analytic methods are used most frequently in the study of nonlinear oscillations because, in comparison with geometric methods, they enable us to get many quantitative results in addition to the qualitative ones. However, these methods are usually efficient only for **weakly nonlinear differential equations**, that is, differential equations differing only slightly from linear differential equations (general nonlinear differential equations are called sometimes **strongly nonlinear differential equations**). In this sense, analytic methods are all \dagger perturbation methods in the wider sense [3], and the variety of the methods lies in the form of perturbation and the method of calculation. To make use of analytic methods, we always reduce the given differential equation to a differential equation of the form

$$\dot{\mathbf{x}} = A\mathbf{x} + \varepsilon \mathbf{X}(\mathbf{x}, t, \varepsilon). \quad (5)$$

Here ε is a parameter with small absolute value, A is a matrix of the form $A = \text{diag}(O_p, B)$, where O_p is a $p \times p$ zero matrix and B is a matrix whose eigenvalues have all nonzero real parts, and $\mathbf{X}(\mathbf{x}, t, \varepsilon)$ is periodic or almost periodic in t .

(i) When $\mathbf{X}(\mathbf{x}, t, \varepsilon)$ is periodic in t , Poincaré's perturbation method is used frequently.

(ii) In practical problems, we frequently meet the case $A = 0$, that is, the case where (5) is of the form

$$\dot{\mathbf{x}} = \varepsilon \mathbf{X}(\mathbf{x}, t, \varepsilon). \quad (6)$$

In this case, the **average** $X_0(\mathbf{x}, \varepsilon) = \lim_{T \rightarrow 0} (1/T) \int_0^T X(\mathbf{x}, t, \varepsilon) dt$ exists. If $dx/dt = X_0(\mathbf{x}, 0)$ has a periodic solution $\xi(t)$ and the related variational linear system is regular, then (6) has an (almost) periodic integral manifold $\mathbf{x} = \mathbf{f}_\varepsilon(t, \theta)$ such that $\mathbf{f}_\varepsilon(t, \theta) \rightarrow \xi(\theta)$ uniformly as $\varepsilon \rightarrow 0$. The **method of averaging** based on this fact was devised by Bogolyubov and Mitropol'skii [4] and is used frequently. The equation

$$\ddot{\mathbf{x}} + \omega^2 \mathbf{x} = \varepsilon f(\mathbf{x}, \dot{\mathbf{x}}, t, \varepsilon) \quad (7)$$

is one of the examples that can be reduced to an equation of the form (6). Equation (7) appears frequently in practical problems, and hence various convenient techniques are devised, such as the **method of linearization**, the **asymptotic method**, the **method of harmonic balance** [4], etc., through which we can apply the method of averaging directly to the given equation (7).

(iii) Consider an ω -periodic system

$$d\mathbf{x}/dt = A(t)\mathbf{x} + \mathbf{f}(\mathbf{x}, t). \quad (8)$$

For any $\mathbf{x}(t) \in P_\omega$, $d\mathbf{x}/dt = A(t)\mathbf{x} + \mathbf{f}(\mathbf{x}(t), t)$ is of the form (4) and it has an ω -periodic solution $\sum_{k=1}^m a_k \xi^k + G[N\mathbf{x}]$, or

$$T\mathbf{x} = \sum_{k=1}^m a_k \xi^k + G \left[N\mathbf{x} - \sum_{k=1}^m (\eta^k, N\mathbf{x}) \eta^k \right] \quad (9)$$

under the condition $(\eta^k, N\mathbf{x}) = 0$ ($k = 1, \dots, m$) as in Section B, where $N: P_\omega \rightarrow P_\omega$ is defined by $N\mathbf{x} = \mathbf{f}(\mathbf{x}(\cdot), \cdot)$, and the ω -periodic solutions of (8) correspond to the fixed points of the transformation $T: P_\omega \rightarrow P_\omega$ induced by (9), where the a_k in (9) are given by $a_k = (\xi^k, \mathbf{x})$, as is required when \mathbf{x} is a fixed point. This is the **alternative or bifurcation method** [6–9], which is extendable to the case of almost periodic systems under the regularity of (3) [10]. In order to obtain a fixed point of T , various kinds of fixed point theorems can be utilized.

(iv) In Section C and (iii) above, the search for a fixed point of an appropriate mapping is a principal device. However, the choice of a suitable domain for the mapping is a crucial problem, and usually an a priori bound for solutions is looked for. The concept of stability and hence Lyapunov's second method are effective in such situations [11].

E. Numerical Methods

Numerical methods are used for obtaining explicit forms of the oscillations. They are convenient in practical applications, since they can be used efficiently whether or not the nonlinearity of the system is weak. For an autonomous or periodic system, one can utilize the following methods as efficient means

of computing periodic solutions: **Newton's iterative method**, the **finite-element method**, the **Lindstedt-Poincaré method**, the **method of multiple scales**, the **method of harmonic balance**, the **Galerkin method**, etc. [12–15].

For weakly nonlinear cases, analytic methods (that is, perturbation methods) are very efficient. However, when a parameter is fixed beforehand, it is not easy to know whether the conclusion obtained by the perturbation method is valid for the given value. For strongly nonlinear cases, it is very difficult to analyze the problems by analytic methods, and at present hardly any efficient methods exist. Numerical methods will therefore become exceedingly important in research on nonlinear oscillations. These are now of tolerable efficiency and reliability, due to progress in high-speed machine computation.

F. Nonstationary Oscillations

Physically speaking, an oscillation is a stationary state. However, when a system contains a parameter varying slowly with time, the oscillation also varies slowly with time (for example, when the length of a pendulum varies slowly, the amplitude of the pendulum also varies slowly). Such variation of oscillations in the course of time is represented by $d\mathbf{x}/dt = \varepsilon f(\mathbf{x}, t, \varepsilon t, \varepsilon)$, where $\mathbf{f}(\mathbf{x}, t, s, \varepsilon)$ is (almost) periodic in t, s . Such cases provide important problems in the theory of nonlinear oscillations; one such problem has been posed by Mitropol'skii [16] under the name **nonstationary oscillations** and has been investigated by means of the method of multiple scales [15].

References

- [1] G. Sansone and R. Conti, *Nonlinear differential equations*, Pergamon, 1964. (Original in Italian, 1956.)
- [2] S. Lefschetz, *Differential equations: Geometric theory*, Interscience, 1959.
- [3] G. E. O. Giacaglia, *Perturbation methods in non-linear systems*, Springer, 1972.
- [4] N. N. Bogolyubov and Yu. A. Mitropol'skii, *Asymptotic methods in the theory of nonlinear oscillations*, Gordon & Breach, 1961. (Original in Russian, 1958.)
- [5] N. Minorsky, *Nonlinear oscillations*, Van Nostrand, 1962.
- [6] J. K. Hale, *Ordinary differential equations*, Wiley, 1969.
- [7] M. A. Krasnosel'skii, *Translation along trajectories of differential equations*, Amer. Math. Soc., 1968. (Original in Russian, 1966.)

[8] L. Cesari, R. Kannan, and J. D. Schuur, *Nonlinear functional analysis and differential equations*, Dekker, 1976.

[9] N. Rouche and J. Mawhin, *Equations différentielles ordinaires II*, Masson, 1973.

[10] M. A. Krasnosel'skii, V. Sh. Burd, and Yu. S. Kolesov, *Nonlinear almost periodic oscillations*, Wiley, 1973. (Original in Russian, 1970.)

[11] T. Yoshizawa, *Stability theory and the existence of periodic solutions and almost periodic solutions*, Springer, 1975.

[12] C. Hayashi, *Nonlinear oscillations in physical systems*, McGraw-Hill, 1964.

[13] A. Andronov, A. Vitt, and S. Khaikin, *Theory of oscillations*, Addison-Wesley, 1966. (Original in Russian, 1959.)

[14] M. Urabe, *Nonlinear autonomous oscillations*, Academic Press, 1967.

[15] A. H. Nayfeh and D. T. Mook, *Nonlinear oscillations*, Wiley, 1979.

[16] Yu. A. Mitropol'skii, *Problems of the asymptotic theory of nonstationary vibrations*, Daniel, 1965. (Original in Russian, 1964.)

291 (XIII.12) Nonlinear Problems

A. General Remarks

Nonlinear problems deal with nonlinear mappings or operators and the related equations. Until recently it was customary to consider nonlinear problems as belonging to applied mathematics and the physical sciences. However, nonlinear problems now belong to modern mathematics. Many phenomena in mathematical physics are essentially described by nonlinear equations, e.g., the motions of several particles or of viscous or compressible fluids (\rightarrow 420 Three-Body Problem, 204 Hydrodynamical Equations). Some of these equations are approximated by linear equations only when the variables appearing in the equations are restricted to very small domains; they are treated by perturbation methods when the variables stay in comparatively small domains. If these requirements cannot be met and we have to deal with equations in which the variation of the variables are not negligible, nonlinear problems certainly arise.

The methods of solution of nonlinear problems are not as powerful or general as those for linear differential equations. For instance, the \dagger principle of superposition of solutions does not hold for nonlinear problems, and therefore Fourier methods are no longer applicable. Indeed, some nonlinear problems can

be dealt with only by means of very particular or ad hoc devices.

B. Methods Used in Nonlinear Problems

Consider a nonlinear equation

$$G(x)=0, \quad (1)$$

where G is a nonlinear mapping or operator of a subset S of a linear space X into itself. If we put $G=I-F$ (I is the identity), equation (1) becomes

$$x=F(x). \quad (2)$$

Then a solution of (2) is a fixed point of F .

Therefore fixed-point theorems of various kinds are useful for solving (2) (\rightarrow 286 Nonlinear Functional Analysis).

Let $x^{(0)} \in S$, and suppose that we can define $x^{(k)}$, $k=1, 2, \dots$, by

$$x^{(k)}=F(x^{(k-1)}), \quad k=1, 2, \dots$$

If F is continuous and the sequence $x^{(k)}$ converges to a point $x \in S$, then x is a fixed point of (2). Such a method of constructing an approximate sequence by iteration is called the **iterative method**. \dagger Newton's iterative process, given by

$$x^{(k)}=x^{(k-1)}-[G'(x^{(k-1)})]^{-1}G(x^{(k-1)}),$$

is one such method, where G' denotes the \dagger Fréchet derivative of G .

If X is an infinite-dimensional space, many concepts and methods of the theory of functional analysis can be used for a number of nonlinear problems (\rightarrow 286 Nonlinear Functional Analysis).

We note that there exist nonlinear transformations that change nonlinear equations into linear ones. For example, the \dagger hodograph method, which is often applied in hydrodynamics, consists of reducing a system of \dagger quasi-linear partial differential equations of the form

$$A_i(u, v)u_x + B_i(u, v)u_y + C_i(u, v)v_x + D_i(u, v)v_y = 0 \quad (i=1, 2)$$

to linear differential equations

$$A_i y_v - B_i x_v - C_i y_u + D_i x_u = 0 \quad (i=1, 2)$$

by means of the \dagger hodograph transformation, which changes the independent variables from x, y to u, v (\rightarrow 205 Hydrodynamics).

C. Nonlinear Algebraic and Transcendental Equations

Consider a system of equations

$$f_i(x_1, \dots, x_n)=0 \quad (i=1, \dots, n). \quad (3)$$

Newton's iterative process can be applied as follows. Starting with a point $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ lying near the desired solution, we define $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$, $k = 1, 2, \dots$, by solving the system of equations

$$\sum_{i=1}^n \frac{\partial f_j}{\partial x_i} (x^{(k-1)}) (x_i^{(k)} - x_i^{(k-1)}) + f_j(x^{(k-1)}) = 0 \quad (j = 1, \dots, n).$$

Under some conditions the iteration $\{x^{(k)}\}$ converges to the solution (\rightarrow 301 Numerical Solution of Algebraic Equations).

If the system (3) is a real one, (3) is equivalent to $\sum f_i^2 = 0$. It is clear that a method of obtaining a minimum of a function $f(x_1, \dots, x_n)$ is applicable to solving $\sum f_i^2 = 0$. Taking a point $x^{(0)}$, we define

$$x^k = x^{(k-1)} - \lambda_{k-1} \nabla f(x^{(k-1)}), \quad k = 1, 2, \dots,$$

where $\nabla f = (\partial f / \partial x_1, \dots, \partial f / \partial x_n)$, and let λ_{k-1} satisfy

$$f(x^{(k-1)} - \lambda_{k-1} \nabla f(x^{(k-1)})) \leq f(x^{(k-1)} - \lambda \nabla f(x^{(k-1)})) \quad (\lambda \geq 0).$$

Under suitable assumptions, a subsequence $x^{(k_j)}$ converges to x such that $\nabla f(x) = 0$ and $f(x^{(k_j)})$ decreases monotonically to $f(x)$ [1].

Let f be a continuous mapping from a domain D of \mathbf{R}^n or \mathbf{C}^n into itself. Then for any $x^{(0)} \in D$, the iteration $x^{(k)}$ can always be defined by $x^{(k)} = f(x^{(k-1)})$. The problem of the behavior of $x^{(k)}$ is an interesting and important one in pure mathematics; recent work in the physical sciences is yielding many new concepts related to this problem (\rightarrow 126 Dynamical Systems, 433 Turbulence and Chaos).

D. Nonlinear Differential Equations

Consider a nonlinear system of ordinary differential equations

$$dx/dt = f(t, x) \quad (x \in \mathbf{R}^n).$$

The initial value problem with initial condition $x(\tau) = \xi$ is equivalent to the problem of solving the nonlinear integral equation

$$x(t) = \xi + \int_{\tau}^t f(s, x(s)) ds. \quad (4)$$

The solution of (4) is a fixed point of the operator $T: \varphi(t) \mapsto \xi + \int_{\tau}^t f(s, \varphi(s)) ds$ defined for a suitable function space. Therefore fixed-point theorems are applicable to T , and the iterative method for T is called the †method of successive approximation. Also the method of the †Cauchy polygon is useful (\rightarrow 316 Ordinary Differential Equations (Initial Value Problems)). These ideas are extended to an

abstract Cauchy problem,

$$du/dt = A(t)u \quad (t > 0), \quad u(+0) = a,$$

in a Banach space X , where $a \in X$ and $A(t)$ is a nonlinear operator (\rightarrow 286 Nonlinear Functional Analysis).

For extensive studies of nonlinear ordinary and partial differential equations \rightarrow 314 Ordinary Differential Equations (Asymptotic Behavior of Solutions), 290 Nonlinear Oscillation, 394 Stability, 321 Partial Differential Equations (Initial Value Problems), 323 Partial Differential Equations of Elliptic Type, 325 Partial Differential Equations of Hyperbolic Type.

Nonlinear differential equations of special types appear in many fields of pure and applied mathematics, e.g., the Monge-Ampère equation and the equation for minimal surfaces in differential geometry (\rightarrow 183 Global Analysis, 275 Minimal Submanifolds), and the Toda lattice equation and the Korteweg-de Vries equation in mathematical physics (\rightarrow 287 Nonlinear Lattice Dynamics, 387 Solitons).

E. Nonlinear Problems of Control Systems

The basic equation for a †control system in which the state of the controlled object can be represented by an n -vector x is given by the following system of differential equations [3]:

$$\dot{x} = Ax - \varphi(\sigma)b, \quad \dot{\xi} = \varphi(\sigma), \quad \sigma = c'x - \gamma\xi, \quad (5)$$

where \dot{x} and $\dot{\xi}$ stand for dx/dt and $d\xi/dt$, respectively, ξ is a scalar function representing the control, b, c are constant n -vectors, γ is a constant number, c' is the transpose of c , and A is a constant $n \times n$ matrix whose characteristic roots have negative real parts. Furthermore, we assume that when the control has no effect upon the system, x is determined by $\dot{x} = Ax$. The quantities under consideration are all real. Finally, the function $\varphi = \varphi(\sigma)$ is a scalar function characteristic of the control mechanism. Generally, φ is nonlinear in σ , and hence equation (5) is nonlinear. Normally we assume that φ has the following properties: (i) $\varphi(\sigma)$ is a real-valued continuous function on $(-\infty, \infty)$ with $\varphi(0) = 0$ and $\sigma\varphi(\sigma) > 0$ for $\sigma \neq 0$; (ii) $\int_0^{\pm\infty} \varphi(\sigma) d\sigma = +\infty$. We say that (5) is **absolutely stable** if

$$x(t) \rightarrow 0, \quad \xi(t) \rightarrow 0 \quad \text{as} \quad t \rightarrow +\infty$$

for any choice of φ subject to (i) and (ii) and for every solution of (5). In the study of control systems an important problem is to obtain a necessary and sufficient condition for the system to be absolutely stable. In this connection, we have the following result, due to M. V.

Popov: (5) is absolutely stable if there exists a nonnegative q such that

$$\operatorname{Re}\{(1+i\omega q)(c'(i\omega I - A)^{-1}b)\} + q\gamma \geq 0 \quad (6)$$

for any real ω , where I is the $n \times n$ identity matrix. Conversely, if the absolute stability of (5) is given by means of the \dagger Lyapunov function (\rightarrow 394 Stability)

$$V(x, \sigma) = x'Bx + \alpha\sigma^2 + \alpha f'x + \beta \int_0^\sigma \varphi(\sigma) d\sigma,$$

where B is a constant matrix and f is a constant vector, then there exists a nonnegative q for which (6) is satisfied for all real ω .

F. Nonlinear Equations in Applied Mathematics

Some examples of nonlinear problems are given here (\rightarrow also 205 Hydrodynamics; 318 Oscillations).

(1) The nonlinear \dagger differential-difference equation

$$du(t)/dt = (a - u(t-1))u(t)$$

is called the **Cherwell-Wright differential equation** [4]. Given the initial condition $u(t) = g(t)$ ($0 \leq t \leq 1$, $g(t)$ is a given continuous function), its solution is uniquely determined for $0 \leq t < \infty$. If $a \leq 0$ and $g(1) > 0$, then $u(t) \rightarrow 0$ as $t \rightarrow \infty$; if $a = 0$ and $g(1) < 0$, then $u(t) \rightarrow -\infty$ as $t \rightarrow \infty$. For $a > 0$, $u(t) \rightarrow -\infty$ as $t \rightarrow \infty$ if $g(1) < 0$, while $u(t)$ either approaches a monotonically ($0 < a \leq 1/e$) or oscillates (boundedly) around a ($a > 1/e$) if $g(1) > 0$. In particular, we have damping oscillations for $a \leq 3/2$, while oscillatory solutions without damping appear for $a > \pi/2$.

(2) If we regard a star as a gas sphere and assume the polytropic relation $p = K\rho^\gamma$ (K and γ are constants) between the pressure p and the density ρ at each point inside the star, then we have a differential equation of the second order that determines the density distribution. This equation constitutes the basis of the classical theory of the internal structure of the stars and is called **Emden's differential equation** or the **polytropic differential equation**. It reads:

$$(1/\xi^2)d(\xi^2 d\theta/d\xi)/d\xi = -\theta^n,$$

where $\gamma = 1 + 1/n$, $\rho = \lambda\theta^n$, $r = \alpha\xi$, $\alpha = ((n+1)K\lambda^{\gamma-2}/(4\pi G))^{1/2}$, with r the distance from the center, G the universal gravitational constant, and λ an arbitrary constant. The solution that satisfies the conditions $\theta = 1$ and $d\theta/d\xi = 1$ at $\xi = 0$ is called the **Lane-Emden function** of index n . Emden's equation is invariant under the transformation $\xi \rightarrow A\xi$, $\theta \rightarrow A^{-2/(n-1)}\theta$ (with A an arbitrary constant).

291 Ref.

Nonlinear Problems

Although Emden's equation can be reduced to an equation of the first order, it has not been possible to solve it analytically except for the cases $n=0, 1$, and 5 . For certain values of n between 0.5 and 6 , Emden gave numerical solutions, which were refined later by Green, D. H. Sadler, and D. C. Miller [5].

(3) **Caianiello's differential equations**, which describe the state of a network of neurons [6], are

$$x_i(t+\tau) = Y\left[\sum_j \sum_r a_{ij}^{(r)} x_j(t-r\tau) - \theta_i\right],$$

where the function $x_i(t)$ represents the state of the i th neuron at the time t and takes only the values 0 and 1 . The state of the system is to be considered at the discrete times $t=0, \tau, 2\tau, \dots$. The (real) coefficient $a_{ij}^{(r)}$ represents the weight of the effect of hysteresis in the relay process from the cell j to the cell i . The nonnegative integer θ_i is the threshold value of the cell i . $Y[x]$ is the unit step function that is equal to 1 for $x \geq 0$ and vanishes for $x < 0$.

(4) The following **Hodgkin-Huxley differential equation** arises in the study of conduction and excitation in nerve systems [7]:

$$\frac{\partial^2 V}{\partial x^2} = \frac{2r_0}{R_0} \left(C_0 \frac{\partial V}{\partial t} + g_1 m^3 h (V - V_1) + g_2 n^4 (V - V_2) + g_3 (V - V_3) \right),$$

$$\frac{\partial m}{\partial t} = -(\alpha_1(V) + \beta_1(V))m + \alpha_1(V),$$

$$\frac{\partial h}{\partial t} = -(\alpha_2(V) + \beta_2(V))h + \alpha_2(V),$$

$$\frac{\partial n}{\partial t} = -(\alpha_3(V) + \beta_3(V))n + \alpha_3(V),$$

where α_i, β_i ($1 \leq i \leq 3$) are given functions of V , and r_0, R_0, C_0, g_i, V_i ($1 \leq i \leq 3$) are constants. The unknown function V is sought in the domain $0 < x < \infty, 0 < t < \infty$, while the initial values of V, m, h, n , and the boundary value of V are given at $t=0$ and $x=0$, respectively.

References

- [1] T. L. Saaty and J. Bram, *Nonlinear mathematics*, McGraw-Hill, 1964 (Dover, 1981).
- [2] T. L. Saaty, *Modern nonlinear equations*, McGraw-Hill, 1967 (Dover, 1981).
- [3] S. Lefschetz, *Stability of nonlinear control systems*, Academic Press, 1965.
- [4] S. Kakutani and L. Markus, On the nonlinear difference-differential equation $y'(t) = (A - By(t-\tau))y(t)$, *Contributions to the theory of nonlinear oscillations*, Ann. Math. Studies, Princeton Univ. Press, 1958, vol. 4, 1-18.

- [5] R. Bellman, Stability theory of differential equations, McGraw-Hill, 1953.
[6] E. R. Caianiello, Outline of a theory of thought-process and thinking machines, J. Theoret. Biol., 1 (1961), 204–235.
[7] A. L. Hodgkin and A. F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, J. Physiol. (London), 117 (1952), 500–544.

292 (XIX.4) Nonlinear Programming

A. Problems

A nonlinear programming problem is a type of mathematical programming problem where it is required to minimize or maximize a nonlinear function $\theta(x)$ of n -vector variable x defined in a closed connected set X^0 with a set of linear or nonlinear constraints. Minimization or maximization of a continuously differentiable function $\theta(x)$ under the equality condition expressed by a set of continuously differentiable functions has been traditionally dealt with by the method of Lagrange multipliers. Hence a typical nonlinear programming problem is usually formulated as follows.

(NLP) Minimize $\theta(x)$ under the condition $\bar{x} \in X^0 \subset \mathbf{R}^n$ and $g_i(x) \leq 0$ for $i = 1, 2, \dots, m$.

Or, equivalently, determine the set of all x such that $\theta(\bar{x}) = \min_{x \in C} \theta(x)$, where $C = \{x | x \in X^0 \text{ and } g(x) \leq 0\}$.

Here we need only consider minimization, since maximization problems can be converted to minimization problems by virtue of the obvious relation $\max \theta = -\min(-\theta)$.

The set C is known as the **feasible region** or the **constraint set**, and \bar{x} is called an **optimal solution** or simply a solution. In many nonlinear programming problems X^0 is \mathbf{R}^n . If $X^0 = \mathbf{R}^n$ and θ and g are linear functions on \mathbf{R}^n , then the problem becomes a linear programming problem (\rightarrow 255 Linear Programming). The problem of minimizing a quadratic function subject to linear constraints is called the **quadratic programming problem** (\rightarrow 349 Quadratic Programming). If X^0 is a convex set, θ is convex (or concave), and the g_i are convex on X^0 , then the minimization (or maximization) problem is known as a **convex** (or **concave**) **programming problem**. Convex and concave functions are important in nonlinear programming, because they admit reasonably straightforward sufficient conditions for optimality and also because they constitute the only important class of functions for which

necessary optimality conditions can be given without the differentiability condition. Subject to suitable modifications, the method of Lagrange multipliers can also be applied to the solution of nonlinear programming problems.

The **Lagrangian function** ψ associated with the minimization problem (NLP) is defined by

$$\psi(x, u) = \theta(x) + u'g(x),$$

where $u = (u_1, \dots, u_m)$ and u' denotes the vector of Lagrange multipliers.

A pair (\bar{x}, \bar{u}) is called a **saddle point** of $\psi(x, u)$, provided $\bar{x} \in X^0$, $\bar{u} \in \mathbf{R}^m$, $\bar{u} \geq 0$, and $\psi(\bar{x}, u) \leq \psi(\bar{x}, \bar{u}) \leq \psi(x, \bar{u})$ for all $x \in X^0$ and all $u \in \mathbf{R}^m$ such that $u \geq 0$. It follows easily that:

(1) (H. Uzawa [18]) If (\bar{x}, \bar{u}) is a saddle point of $\psi(x, u)$, then \bar{x} is an optimal solution.

(2) (Kuhn and Tucker [11]) Assume that X^0 is open and convex, and that θ and g are differentiable and convex. If there exists a pair (\bar{x}, \bar{u}) such that

$$\nabla \theta(\bar{x}) + \bar{u}' \nabla g(\bar{x}) = 0, \quad \bar{x} \in C, \quad \bar{u}' g(\bar{x}) = 0, \quad \bar{u} \geq 0$$

(where $\nabla \theta(x)$ and $\nabla g(x)$ denote, respectively, the gradient vector of θ at x and the Jacobian matrix of g at x), then \bar{x} is an optimal solution.

B. Necessary Conditions for Optimality

In Section A sufficient conditions for optimality were given. Some necessary conditions are also known.

(3) When $\theta(x)$ is continuously differentiable and all the g_i are linear, that is, when the condition is expressed as $Ax \leq b$ and $x \geq 0$, where A is an $m \times n$ matrix and b an m -vector, \bar{x} is an optimal solution only if \bar{x} is an optimal solution of the following linear programming problem.

(LP⁰) Minimize $v'x$ under the condition that $Ax \leq b$, $x \geq 0$, where $v = \nabla \theta(\bar{x})$.

By applying the duality theorem of linear programming, it can be proved that there exists a vector $\bar{u} \geq 0$ such that

$$\nabla \varphi(\bar{x}, \bar{u}) = \nabla \theta(\bar{x}) + \bar{u}' A = 0.$$

(4) (F. John [9]) Assume that X^0 is open, and that θ and g are differentiable. Then, if \bar{x} is a solution of the problem (NLP), there exist $u_0 \geq 0$ and $\bar{u} \in \mathbf{R}^m$ such that

$$u_0 \nabla \theta(\bar{x}) + \bar{u}' \nabla g(\bar{x}) = 0,$$

$$g(\bar{x}) \leq 0, \quad \bar{u}' g(\bar{x}) = 0, \quad \bar{u} \geq 0.$$

Now we consider the following: The vector-valued function g is said to satisfy **Guignard's constraint qualification** at an inner point x of X^0 if any vector y satisfying the linear inequalities $\nabla g_i(x) \cdot y \leq 0$ for $i \in I = \{i | g_i(x) = 0\}$ is in the convex hull spanned by the vectors tangent to the set C at x .

(5) (Guignard [7]) Let X^0 be an open set, let \bar{x} be an optimal solution of (NLP), let θ and g be differentiable at \bar{x} , and assume that g satisfies the Guignard constraint qualification. Then there exists $\bar{u} \in \mathbf{R}^m$ such that $\nabla \theta(\bar{x}) + \bar{u}' \nabla g(\bar{x}) = 0$, $g(x) \leq 0$, and $\bar{u}' g(\bar{x}) = 0$, $\bar{u} \geq 0$.

Guignard's constraint qualification is satisfied if neither of the following conditions hold: (L) Vectors $g_i(x)$ for $i \in I$ are linearly independent.

(S) (Slater's constraint qualification) The g_i are convex and X^0 is a convex set; and there exists a vector x such that $g_i(x) < 0$ for all i . With convexity of θ and g_i , differentiability is not required:

(6) (Kuhn and Tucker [11]) Let X^0 be a convex set, let θ and g be convex on X^0 , and assume that g satisfies Slater's constraint qualification. If \bar{x} is an optimal solution, then there exists $\bar{u} \in \mathbf{R}^m$, $\bar{u} \geq 0$, such that $\bar{u}' g(\bar{x}) = 0$ and (\bar{x}, \bar{u}) is a saddle point of $\psi(x, u) = \theta(x) + u' g(x)$.

C. Sensitivity Analysis

Now consider the following class of problems.

(Pc) Minimize the function $\theta(x)$ under the condition that $g(x) \leq c$, where c is a real m -vector. We denote by \bar{x}_c the set of the solutions of the problem (Pc) and denote $\theta_c^* = \theta(\bar{x}_c)$, $\bar{x}_c \in \bar{X}_c$. Suppose that Guignard's condition is satisfied for each c and that the set of Lagrange multipliers Λ_c is nonempty. Then for any real vector a , we have

$$\inf a' u_c \leq \lim_{t \rightarrow 0} \frac{1}{t} (\theta_{c+ta}^* - \theta_c^*) \leq \sup a' u_c, \quad u_c \in \Lambda_c.$$

Therefore, if the Lagrange multiplier is uniquely determined for some c , then u_c represents the vector of the rates of increments of the objective function to the small increments of the components of the constraints vector. Hence the components of the Lagrange-multiplier vector are called the **imputed prices** or **shadow costs** of the constraints; these have important economic implications, especially when the objective function is expressed in terms of money or profits.

More generally, we can consider the following class of problems.

(GPc): Minimize the function $\theta(x, c)$ under the condition that $g(x, c) \leq 0$, where c is a real parameter.

Denote by $\bar{x}(c)$ and $u(c)$ the solution and the Lagrange multiplier of (GPc) corresponding to the parameter c , respectively, and let $\theta^*(c) = \theta(\bar{x}(c))$. Then under a set of regularity conditions we have

$$\nabla \theta^*(c) = \nabla_c \theta(x(c), c) + u(c)' \nabla_x g(x(c), c),$$

where ∇_c denotes the gradient of θ or g with respect to the parameter.

D. Duality

A **duality theorem** in mathematical programming is the statement of a certain relationship between two problems. This relationship has the following two aspects: (i) one problem is a constrained minimization problem and the other is a constrained maximization problem; (ii) the existence of a solution to one of these problems ensures the existence of a solution to the other, and in this case their respective values are equal.

Let $\psi(x, u)$ be the Lagrangian form of (NLP), and define $\omega(u) = \inf_{x \in X^0} \psi(x, u)$. Then we can state the following two problems.

(P) (primary problem) Minimize $\theta(x)$ under the condition that $x \in X^0$ and $g(x) \leq 0$

(D) (dual problem) Maximize $\omega(u)$ under the condition $u \geq 0$.

If (\bar{x}, \bar{u}) gives a saddle point of $\psi(x, u)$, then

$$\inf \theta(x) = \theta(\bar{x}) = \sup \omega(u) = \omega(\bar{u}) = \psi(\bar{x}, \bar{u}).$$

The dual problem can be formulated alternatively as follows.

(\tilde{D}) Maximize $\psi(x, u) = \theta(x) + u' g(x)$ subject to $(x, u) \in Y = \{(x, u) | x \in X^0, u \in \mathbf{R}^m, u \geq 0, \nabla_x \psi(x, u) = 0\}$ (where $\nabla_x \psi(x, u)$ denotes the vector whose components are the partial derivatives $\partial \psi(x, u) / \partial x_i$ for $i = 1, \dots, n$).

There are a number of duality theorems related to problems (P) and (\tilde{D}); two such theorems are as follows:

(1) (P. Wolfe [20]) Suppose that X^0 is open and convex, that θ and g are differentiable and convex, and that g satisfies the Kuhn-Tucker constraint qualification. Then, if \bar{x} is a solution of (P), there exists a $\bar{u} \in \mathbf{R}^m$ such that (\bar{x}, \bar{u}) is a solution of (\tilde{D}) and $\theta(\bar{x}) = \psi(\bar{x}, \bar{u})$.

(2) (O. L. Mangasarian and J. Ponsstein [13]) Suppose that X^0 is open and convex, that θ and g are differentiable and convex, and that (\hat{x}, \hat{u}) is a solution of (\tilde{D}). If $\psi(x, u)$ is strictly convex in some neighborhood of \hat{x} , then \hat{x} is a solution of (P) and $\theta(\hat{x}) = \psi(\hat{x}, \hat{u})$.

The above two problems (P) and (D) are not symmetric. The notion of symmetric duality was introduced by G. B. Dantzig, E. Eisenberg, and R. W. Cottle:

Primary: Minimize

$$F(x, u) = K(x, u) - u' \nabla_u K(x, u),$$

subject to the constraints

$$\nabla_u K(x, u) \leq 0, \quad x \geq 0, \quad \text{and } u \geq 0;$$

Dual: Maximize

$$G(x, u) = K(x, u) - x' \nabla_x K(x, u),$$

subject to the constraints

$$\nabla_x K(x, u) \geq 0, \quad x \geq 0, \quad \text{and } u \geq 0,$$

where K is continuously differentiable in $(x, u) \in \mathbf{R}^n \times \mathbf{R}^m$.

(3) Dantzig et al. proved [6] the existence of a common optimal solution (\bar{x}, \bar{u}) to both the primary and dual problems, provided (i) an optimal solution (\bar{x}, \bar{u}) to the primary problem exists, (ii) K is convex in x for each u and concave in u for each x , and (iii) K is twice differentiable and the matrix of second partials $(\partial^2 K / \partial u^i \partial u^j)$ is negative definite at (\bar{x}, \bar{u}) .

Rockafellar [15] gave another expression of the duality relation: Define $F(x, y) = \theta(x)$ if $g(x) \leq y$ and $= \infty$ otherwise, and denote $\varphi(y) = \inf_{x \in X^0} F(x, y)$. Then $\theta(x) = \varphi(0)$. For any nonlinear function $\varphi(y)$, the conjugate $\varphi^*(\eta)$ is defined by

$$\varphi^*(\eta) = \sup_y (\eta'y - \varphi(y)).$$

(4) (Rockafellar [15]) $\sup \omega(u) = \varphi^{**}(0) = \text{clco } \varphi(0)$, where $\text{clco } \varphi(y)$ denotes the closed convex hull or maximum convex minorant of $\varphi(y)$ defined by $\text{clco } \varphi(y) = \sup_i 1'y \{1'z \leq \varphi(z) \text{ for all } z\}$. It follows from the above that $\inf \theta(x) = \sup \omega(u)$ if and only if $\varphi(0) = \text{clco } \varphi(0)$, which holds true if $\varphi(y)$ is convex.

Further forms of the duality theorem hold for linear or quadratic programming problems (\rightarrow 255 Linear Programming, 349 Quadratic Programming).

E. Algorithms

In a limited class of problems, i.e., when the objective function is quadratic and the constraints are linear (\rightarrow 349 Quadratic Programming) the optimal solution can be obtained by solving a system of linear equations by the simplex method or other algorithms; but in most nonlinear programming the solution is calculated by some kind of iterative procedure. Note that even when the constraints are given as equalities and all the functions involved are continuously differentiable, so that the optimal solution is explicitly given as a solution of a set of simultaneous equations, we usually require some iteration procedure, such as the Newton-Raphson algorithm, to obtain numerically the solution with preassigned accuracy; and the iteration is not always easy if the functions are sufficiently complex.

Several iterative procedures for solving nonlinear programming problems have been proposed. Since the simplex method is a powerful tool in linear programming, one type of approach is to obtain an approximately optimal solution by approximating the objective and the constraint functions by piecewise linear functions and then applying linear programming techniques to get the approximately optimal solution within each region. Another

is to change the constrained problem to a nonconstrained one by introducing a sufficiently large number M , called the penalty, and then maximizing

$$\theta_M(x) = \theta(x) + M \sum_j \max(g_j(x), 0)$$

without the constraint. This is called the **penalty method**.

The third and most generally applicable technique is the **gradient method**, of which several variations are known:

(i) **Arrow-Hurwicz-Uzawa gradient method** [4]. Concave or convex programming problems can be solved by finding a saddle point of the Lagrangian function $\psi(x, u)$. Let $\varphi(x, u)$ be strictly concave and of class C^2 in n -vector $x \geq 0$ and convex and of class C^2 in m -vector $u \geq 0$ and possess a saddle point (\bar{x}, \bar{u}) . To approach a saddle point of $\varphi(x, u)$ it is natural to devise a gradient process of the form

$$\frac{dx_i}{dt} = \frac{\partial \varphi}{\partial x_i}, \quad \frac{du_j}{dt} = -\frac{\partial \varphi}{\partial u_j}. \quad (1)$$

To keep the variables in the positive orthant, we need to modify (1), and we consider the following system of differential equations:

$$\frac{dx_i}{dt} = \begin{cases} 0 & \text{if } x_i = 0 \text{ and } \frac{\partial \varphi}{\partial x_i} < 0 \\ \frac{\partial \varphi}{\partial x_i} & \text{otherwise} \end{cases} \quad (i = 1, \dots, n),$$

$$\frac{du_j}{dt} = \begin{cases} 0 & \text{if } u_j = 0 \text{ and } \frac{\partial \varphi}{\partial u_j} > 0 \\ -\frac{\partial \varphi}{\partial u_j} & \text{otherwise} \end{cases} \quad (j = 1, \dots, m).$$

Under certain regularity hypotheses, there exists a unique solution $(x(t), u(t))$ of the system with any initial point (x^0, u^0) , and the x component $x(t)$ of the solution converges to \bar{x} as $t \rightarrow \infty$.

Applying the above results to the Lagrangian function $\psi(x, u)$ we can solve the concave or convex programming problem.

(ii) **Rosen's gradient projection method** [16]. If a point x^0 of the feasible region does not give a solution for the minimization problem (P), then we look for a feasible point with a lower function value by proceeding from x^0 in the direction of the gradient of the function $-\theta(x)$. The method fails if x^0 is a boundary point and if the gradient vector points toward the exterior of the feasible region. Rosen's method [16] is to project the gradient onto the boundary of the feasible region and then proceed in the direction of this projection. In this manner, we remain on the boundary of the feasible region.

(iii) **Methods of feasible directions.** These were first described by G. Zoutendijk [21].

Consider the problem of minimizing $\theta(x)$ subject to the constraint $x \in S \subset \mathbb{R}^n$, where S is a closed, connected set satisfying certain regularity conditions and θ is a continuously differentiable function of the n -vector x , such that, for some α , the set $\{x \in S \mid \theta(x) \leq \alpha\}$ is bounded and nonempty.

A method of feasible directions is any recipe for solving this problem by proceeding along the following lines: (1) Start with some $x^0 \in S$ such that $\theta(x^0) < \sup \alpha$. (2) Pass from the k th iteration point x^k to x^{k+1} by first determining a direction s^k in x^k such that the ray $x^k + \lambda s^k$ lies in S for all sufficiently small $\lambda > 0$. (3) Then determine the step length λ_k , thus obtaining the $(k+1)$ st iterate $x^{k+1} = x^k + \lambda_k s^k$. (4) Repeat this procedure until some prescribed stopping condition is satisfied.

There are many methods of determining the s^k , and in most cases the λ_k are then determined by solving a one-dimensional minimum problem along the direction so obtained. Zoutendijk has unified the various possible methods of feasible directions and the relevant normalization rules that yield the optimal directions s^k , and has investigated this subject in detail from the viewpoint of computational technique [21, 22].

F. Generalizations

The extension of the Kuhn-Tucker theory to linear topological spaces is due to L. Hurwicz [8]. Let \mathcal{X} be a linear space, \mathcal{Y}, \mathcal{Z} be linear topological spaces, P_Y, P_Z the nonnegativity cones of \mathcal{Y}, \mathcal{Z} , respectively, which are closed convex cones containing inner points, D a convex set in \mathcal{X} , and F, G concave mappings (\rightarrow 88 Convex Analysis A) from D into \mathcal{Y}, \mathcal{Z} , respectively, such that $G(D)$ contains an inner point of P_Z . If $F(X)$ attains its maximal point when $X = X_0$ and X_0 satisfies $G(X_0) \geq 0, X_0 \in D$, then there exist $Y_0^* \geq 0, Z_0^* \geq 0$ such that $\Phi(X, Z^*) = Y_0^*(F(X)) + Z_0^*(G(X))$ has a saddle point at (X_0, Z_0^*) . The condition that P_Y and P_Z have inner points can be weakened to cover the cases of $(l_p), (L_p), (s), (S)$ (Hurwicz and Uzawa). P. P. Varaiya [19] considered the following nonlinear programming problem in Banach space.

(B) Maximize $f(x)$ subject to $x \in A, g(x) \in A_Y$, where X, Y are real Banach spaces, $x \in X, g: X \rightarrow Y$ is a Fréchet differentiable mapping, f is a real-valued differentiable function, A is a subset of X , and A_Y is a convex set in Y .

The main results are similar to the Kuhn-Tucker necessary conditions. Varaiya also exhibited a saddle value problem related to

(B), when A_Y is a closed convex cone. L. W. Neustadt [14] investigated nonlinear programming problems in linear vector spaces and gave an application to the theory of optimal control. The main results are: (i) Kuhn-Tucker type conditions which are both necessary and sufficient for optimality, (ii) a duality theory for obtaining multipliers in the generalized Kuhn-Tucker conditions, and (iii) an application to optimal control theory.

References

- [1] J. Abadie (ed.), *Nonlinear programming*, North-Holland, 1967.
- [2] J. Abadie (ed.), *Integer and nonlinear programming*, North-Holland, 1970.
- [3] K. J. Arrow, L. Hurwicz, and H. Uzawa (eds.), *Studies in linear and nonlinear programming*, Stanford Univ. Press, 1958.
- [4] K. J. Arrow and L. Hurwicz, Gradient method for concave programming I, in [3, pp. 117–126]; H. Uzawa, Gradient method for concave programming II, in [3, pp. 127–132].
- [5] E. M. Beale, On minimizing a convex function subject to linear inequalities, *J. Roy. Statist. Soc., (B)* 17 (1955), 173–184.
- [6] G. B. Dantzig, E. Eisenberg, and R. W. Cottle, Symmetric dual nonlinear programs, *Pacific J. Math.*, 15 (1965), 809–812.
- [7] M. Guignard, Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space, *SIAM J. Control*, 7 (1969), 232–241.
- [8] L. Hurwicz, Programming in linear spaces, in [3, pp. 38–102].
- [9] F. John, Extremum problems with inequalities with subsidiary conditions, *Studies and Essays: Courant Anniversary Volume*, Interscience, 1948, 187–204.
- [10] H. P. Künzi and W. Krelle, *Nichtlineare Programmierung*, Springer, 1962.
- [11] H. W. Kuhn and A. W. Tucker, Nonlinear programming, *Proc. 2nd Berkeley Symp. Math. Statist. Prob.*, 1951, 481–492.
- [12] O. L. Mangasarian, *Nonlinear programming*, McGraw-Hill, 1969.
- [13] O. L. Mangasarian and J. Ponstein, Minimax and duality in nonlinear programming, *J. Math. Anal. Appl.*, 11 (1965), 504–518.
- [14] L. W. Neustadt, Sufficiency conditions and a duality theory for mathematical programming problems in arbitrary linear spaces, in [17, pp. 323–348].
- [15] R. T. Rockafellar, Augmented Lagrange multiplier functions and duality in nonconvex programming, *SIAM J. Control*, 12 (1974), 293–322.
- [16] J. B. Rosen, The gradient projection method for nonlinear programming, *J. Soc.*

- Indust. Appl. Math., 8 (1960), 181–217; 9 (1961), 514–553.
 [17] J. B. Rosen, O. L. Mangasarian, and K. Ritter (eds.), *Nonlinear programming*, Academic Press, 1970.
 [18] H. Uzawa, The Kuhn-Tucker theorem in concave programming, in [3, pp. 32–37].
 [19] P. P. Varaiya, Nonlinear programming in Banach space, *SIAM J. Appl. Math.*, 15 (1967), 284–293.
 [20] P. Wolfe, A duality theorem for nonlinear programming, *Quart. Appl. Math.*, 19 (1961), 239–244.
 [21] G. Zoutendijk, *Methods of feasible directions*, Elsevier, 1960.
 [22] G. Zoutendijk, Nonlinear programming, computational methods, in [2, pp. 37–86].
 [23] G. Zoutendijk, Some algorithms based on the principle of feasible directions, in [17, pp. 93–121].

293 (I.7) Nonstandard Analysis

A. General Remarks

Nonstandard analysis is a new field of research that has branched off from model theory and that provides a powerful method applicable in almost all fields of mathematical science.

About 1960, A. Robinson successfully used a nonstandard model of the real number field \mathbf{R} to justify Leibniz-type infinitesimal calculus. After this, using higher-order logic, he developed a stronger theory of nonstandard analysis, and applied it to other mathematical fields [1].

In this article, we adopt a first-order logic over a universe. In the final section we present a theory called nonstandard set theory; this is a conservative extension of Zermelo-Fraenkel set theory with the axiom of choice.

B. Axioms for Nonstandard Analysis

A nonempty set U is called a universe if the following four conditions are satisfied:

- (a) $x \in U, y \in x \Rightarrow y \in U$;
- (b) $x, y \in U \Rightarrow \{x, y\} \in U$;
- (c) $x \in U \Rightarrow \bigcup x \in U$;
- (d) $x \in U \Rightarrow \mathbf{P}(x) \in U$.

In what follows, U will be a universe containing the field \mathbf{R} of real numbers.

We construct a language \mathcal{L} describing mathematics in U . The alphabet consists of the

following symbols: (1) countably many variables. (2) constants corresponding to all elements in U . (3) predicate symbols $=$ and \in . (4) logical symbols $\neg, \wedge, \vee, \rightarrow, \exists, \forall$. (5) auxiliary symbols $[,]$. The last two symbols will be omitted where there is no danger of confusion.

Definition of formulas. (1) If t and s are terms (variables or constants), then $t = s$ and $t \in s$ are formulas (atomic formulas). (2) If ϕ and ψ are formulas, then $\neg\phi, \phi \wedge \psi, \phi \vee \psi, \phi \rightarrow \psi$ are formulas. (3) If ϕ is a formula and x is a variable, then $\exists x[\phi]$ and $\forall x[\phi]$ are formulas. (4) The formulas are those that can be constructed by the above procedure.

A formula ϕ containing at most n free variables is called an n -ary formula. In this case, we often write $\phi(x_1, \dots, x_n)$ for ϕ . A sentence is a 0-ary formula. We write $\models \phi$ if a sentence ϕ is true under the usual interpretation.

Consider a quadruple $(U, *U, *e, *)$ where $*U$ is a set, $*e$ is a binary relation on $*U$, and $*$ is a mapping: $a \mapsto *a$ from U to $*U$. An element α of $*U$ is called **standard** if $\alpha = *a$ for some $a \in U$, and **nonstandard** if not. Adjoin to \mathcal{L} constants representing all nonstandard elements of $*U$. Then we have a language $*\mathcal{L}$ for $*U$. A sentence in \mathcal{L} is a sentence in $*\mathcal{L}$. Taking $*U$ as the scope of quantifiers, we can interpret every $*\mathcal{L}$ -sentence in $*U$. We write $*\models \phi$ if an $*\mathcal{L}$ -sentence ϕ is true under this interpretation.

Axiom 1 (transfer): For every sentence ϕ in \mathcal{L} , we have

$$\models \phi \Leftrightarrow *\models \phi.$$

Definition: Let $\phi(x, y)$ be a binary formula in \mathcal{L} (resp. in $*\mathcal{L}$). We say that $\phi(x, y)$ is **concurrent** in U (resp. in $*U$), if, for every finite number of elements a_1, \dots, a_n in U (resp. in $*U$), there exists an element b in U (resp. in $*U$) such that $\models \phi(a_i, b)$ (resp. $*\models \phi(a_i, b)$) for $1 \leq i \leq n$.

Axiom 2 (enlargement): If a binary formula $\phi(x, y)$ in \mathcal{L} is concurrent in U , there exists an element β in $*U$ such that $*\models \phi(a, \beta)$ for all a in U .

A quadruple $(U, *U, *e, *)$ or simply $*U$ is called an **enlargement** of U if it satisfies Axioms 1 and 2.

These axioms are strong enough to develop the basic theory of nonstandard analysis, but sometimes we require a stronger axiom. Let κ be an infinite cardinal.

Axiom 3 (κ -saturation): Let $\phi(x, y)$ be a binary $*\mathcal{L}$ -formula concurrent in $*U$. Then, for every subset A of $*U$ with cardinality at most κ , there exists an element β of $*U$ such that $*\models \phi(\alpha, \beta)$ for all $\alpha \in A$.

$*U$ is called a **κ -saturated model** if it satisfies Axioms 1 and 3, and a **κ -saturated enlarge-**

ment if it satisfies Axioms 1, 2, and 3. If κ is not less than the cardinality of U , then a κ -saturated model is necessarily an enlargement.

In the remainder of this article, *U will be an enlargement of U , if the contrary is not explicitly mentioned.

For α in *U , $\hat{\alpha}$ is the set of $\xi \in ^*U$ such that $\xi^* \in \alpha: \hat{\alpha} = \{\xi \in ^*U \mid \xi^* \in \alpha\}$. For simplicity, we write \in for $^*\in$ and identify $\hat{\alpha}$ with α . Under this identification, a subset A of *U is called **internal** if it belongs to *U and **external** if not.

Axiom 1 and 2 imply the following results.

- (1) There exist infinite hypernatural numbers, namely, elements of $^*\mathbb{N}$ bigger than all $^*n, n \in \mathbb{N}$.
- (2) There are positive infinitesimal hyperreal numbers.
- (3) For every set A in U , there exists a hyperfinite internal subset of *A containing all $^*a, a \in A$. Here, Γ is called **hyperfinite** if we have $^*\models \phi(\Gamma)$, where $\phi(x)$ is a formula that says “ x is a finite set.”

- (4) Let A be a set in U . Then *A has non-standard elements if and only if A is an infinite set.

- (5) If a family \mathcal{F} of sets in U has the finite intersection property, then the family $\{^*A \mid A \in \mathcal{F}\}$ in *U has a nonempty intersection.

- (6) Let $\alpha: ^*\mathbb{N} \rightarrow ^*\mathbb{R}$ be an internal hypersequence. If $\alpha(^*n)$ is infinitesimal for every $n \in \mathbb{N}$, then there exists an infinite hypernatural number λ such that $\alpha(v)$ is infinitesimal for every hypernatural number v less than λ .

If we moreover assume Axiom 3, we have the following results.

- (7) The cardinality of an internal set is either finite or more than κ .

- (8) If a family \mathcal{F} of internal sets in *U has the finite intersection property and cardinality κ at most, \mathcal{F} has a nonempty intersection.

- (9) Introduce the order topology in $^*\mathbb{R}$. Then every subset of $^*\mathbb{R}$ with cardinality at most κ is bounded and discrete.

- (10) Let α, β be internal sets and C a subset of α with cardinality at most κ . Then every mapping from C to β can be prolonged to an internal mapping from α to β .

- (10') In particular, every external sequence $\mathbb{N} \rightarrow \beta$ can be prolonged to an internal sequence $^*\mathbb{N} \rightarrow \beta$. This follows from the assumption of countable saturation. This fact plays an essential role in nonstandard probability theory, which is now in the process of rapid development (\rightarrow Section D (3)).

C. Construction of Ultrapower Models

Let I be an infinite set. A mapping α from I to U is a family of elements in U with indices in I . So we write $\langle \alpha(i) \rangle_{i \in I}$ or simply $\langle \alpha(i) \rangle$ for α .

Let \mathcal{F} be a nonprincipal ultrafilter on I and define an equivalence relation $\sim_{\mathcal{F}}$ on the set U^I of all mappings from I to U :

$$\langle \alpha(i) \rangle \sim_{\mathcal{F}} \langle \beta(i) \rangle \Leftrightarrow \{i \in I \mid \alpha(i) = \beta(i)\} \in \mathcal{F}.$$

Denote by *U the quotient set of U under the relation $\sim_{\mathcal{F}}$. The class of $\langle \alpha(i) \rangle_{i \in I}$ will be denoted by $[\alpha(i)]_{i \in I}$ or simply $[\alpha(i)]$.

Define a binary relation $^*\in$ on *U :

$$[\alpha(i)]^* \in [\beta(i)] \Leftrightarrow \{i \in I \mid \alpha(i) \in \beta(i)\} \in \mathcal{F}.$$

Let $*$ be the diagonal mapping from U to *U and consider the quadruple $(U, ^*U, ^*\in, *)$.

Theorem (Łoś): The foregoing quadruple satisfies Axiom 1. Moreover, *U is a countably saturated model.

In particular, let I be the set of all finite subsets of U . For $i \in I$, put $\mu(i) = \{j \in I \mid i \subset j\}$. Then the family of sets $\mathcal{B} = \{\mu(i) \mid i \in I\}$ has the finite intersection property, and therefore there exists an ultrafilter \mathcal{F} on I including \mathcal{B} (use $^*\text{Zorn's lemma}$). If we construct *U from the ultrafilter \mathcal{F} , then *U is a countably saturated enlargement of U [1].

It is not easy to construct a κ -saturated enlargement for an arbitrary cardinal κ . We have two ways: to use a κ -good ultrafilter (the proof of its existence is difficult) [2, 3] or to use an ultralimit (iteration of ultrapowers) [4, 16].

D. Applications

(1) Infinitesimal Calculus. A hyperreal number α (element of $^*\mathbb{R}$) is called **infinitesimal** if $|\alpha|$ is smaller than every positive real number. If $\alpha - \beta$ is infinitesimal, α and β are said to be infinitely close to each other; this is written $\alpha \approx \beta$. If $1/\alpha$ is not infinitesimal, α is called finite. Every finite hyperreal number α is infinitely close to a real number a (completeness of \mathbb{R}). We call a the **standard part** of α and write $\text{st}(\alpha)$. Every finite hyperinteger is an integer.

Let f be a real-valued function on a real interval I . Then *f is a hyperreal-valued function on the hyperreal interval *I . The function f is continuous if and only if $f(x) \approx ^*f(\eta)$ for every $x \in I, \eta \in ^*I$ with $x \approx \eta$. The function f is uniformly continuous if $^*f(\xi) \approx ^*f(\eta)$ for every $\xi, \eta \in ^*I$ with $\xi \approx \eta$.

Let δx be a variable ranging over nonzero infinitesimals. The function $\delta f = ^*f(a + \delta x) - f(a)$ of δx will be called the infinitesimal increment of f at a . f is differentiable at a if and only if the quotient $\delta f / \delta x$ is of infinitesimal variation. The common standard part of $\delta f / \delta x$ is the derivative $f'(a)$.

The higher-order differential $\delta^n f$ is also justified as a higher-order infinitesimal difference, with the standard part of $\delta^n f / \delta x^n$ being

$f^{(n)}(a)$. We can define the Riemann integral as the standard part of a Riemann hypersum with respect to a hyperfinite partition of infinitesimal width.

This type of reformulation permits us to rewrite the whole of calculus, and many teachers are trying to adapt this theory to elementary calculus [5, 6].

(2) Topological Spaces. Let X be a topological space in U and let a be a point of X . The intersection of $*A$, A varying over the neighborhoods of a is called the **monad** of a and is denoted by $\text{Mon}(a)$. There exist hyperneighborhoods of a contained in $\text{Mon}(a)$ (infinitesimal neighborhoods). The topology is determined by the system of monads $\langle \text{Mon}(a) \rangle_{a \in X}$. We can thus rewrite the theory of topological spaces. For example, X is Hausdorff if and only if $\text{Mon}(a) \cap \text{Mon}(b) = \emptyset$ for every $a \neq b$ in X . X is compact if and only if every point of $*X$ belongs to the monad of some element in X . This characterization is very useful; using it, Robinson and Bernstein were able to solve the invariant subspace problem for a special class of operators on a Hilbert space [7]. We can also construct a Haar measure very naturally and simply [2].

(3) Measures and Probability Theory. Let (X, \mathbf{B}, m) be a measure space. Then there exist a hyperfinite subset Γ of $*X$ and a positive hyperreal-valued internal function φ such that we have

$$\int_X f \, dm = \sum_{\xi \in \Gamma} {}^*f(\xi) \varphi(\xi)$$

for every integrable function f [8]. The right-hand side is a hyperfinite sum. It follows that the measure m can be extended to a finitely additive measure defined over all subsets of X .

If in particular every measurable finite set is of measure 0, then there exist a hyperfinite subset Γ of $*X$ including X and a hypernatural number ρ such that we have

$$\int_X f \, dm \approx \frac{1}{\rho} \sum_{\xi \in \Gamma} {}^*f(\xi)$$

for every integrable function f [8]. If $m(X) = 1$, then ρ can be taken as the hypercardinality of Γ . Here, the right-hand side is nothing but the mean value of $*f$ on a hyperfinite set. This idea leads us to a simple description of probability theory.

The above method is rather formal; but P. Loeb [9] has pioneered a new approach in probability theory by constructing an external measure space from a hyperfinitely additive internal measure space (X, \mathcal{A}, ν) in $*U$. In the following, $*U$ is supposed to be countably saturated.

Let $\sigma(\mathcal{A})$ be the external countably additive algebra of subsets of X generated by \mathcal{A} . Then the mapping $\mathcal{A} \mapsto \text{st}(\nu(\mathcal{A}))$ from \mathcal{A} to \mathbf{R} can be extended to a measure on $\sigma(\mathcal{A})$. The completion of this measure space is denoted $(X, L(\mathcal{A}), L(\nu))$. In cases where X is hyperfinite, \mathcal{A} is the totality of internal subsets of X , and $\nu(X) = 1$, then the space $(X, L(\mathcal{A}), L(\nu))$ is called a **Loeb space** and $L(\nu)$ a **Loeb measure**.

Every Radon probability space can be represented as the image of a Loeb space by a measure-preserving mapping. Therefore the probability theory on a Radon probability space can be reduced to that on a Loeb space.

For example, for t Lebesgue measure on the interval $[0, 1]$, take an infinite hypernatural number λ and put $X = \{\mu/\lambda \mid \mu \in {}^*\mathbf{N}, 0 \leq \mu \leq \lambda - 1\}$. If we assign $1/\lambda$ to every point of X , we have an internal probability measure ν on X . Then the mapping $\text{st}: x \mapsto \text{st}(x)$ from X to $[0, 1]$ serves as a measure-preserving mapping from the Loeb space $(X, L(\mathcal{A}), L(\nu))$ onto the Lebesgue measure space on $[0, 1]$.

The notion of **lifting** plays a key role in probability theory on Loeb spaces. Let f be a real-valued function on X , and F an internal hyperreal-valued function on X . F is called a lifting of f if we have $f(x) = \text{st}(F(x))$ almost everywhere with respect to $L(\nu)$. Hence f is measurable if and only if it has a lifting.

Shuttling between an internal probability space and its Loeb space by lifting and standard-part mapping, we can develop, simply and partially hyperfinitely, probability theory on the Loeb space. Among others, Anderson [10, 11] and Keisler [12] applied this method quite successfully to Brownian motions, Itô integrals, stochastic differential equations, etc. [17].

E. Nonstandard Set Theory

In our formulation in \mathbf{B} , we must construct $*U$ from a fixed universe U . Nelson [13], Hrbacek [14] and Čuda [15] independently invented theories that nonstandardize the whole of set theory. In these theories, there is only one real number field \mathbf{R} , and \mathbf{R} already contains infinitesimal numbers. Compare this to Robinson-type infinitesimal analysis, where infinitesimal numbers are introduced as an ad hoc tool. The new theories may demand a reflection on mathematical description in the natural sciences.

In the following, we outline (intuitively) Hrbacek's theory, as strengthened and improved by Kawai [16].

We start from ZFC, the Zermelo-Fraenkel set theory plus the axiom of choice. The language of Kawai's theory NST is that of ZFC

plus two constants S and I . We understand S as the totality of standard sets and I as the totality of internal sets.

Let ϕ be a formula of ZFC, that is, a formula of NST without S and I . We write $^S\phi$ (resp. $^I\phi$), the formula of NST obtained by restricting the scope of variables in ϕ to S (resp. I).

The mathematical axioms and axiom schemes of NST are substantially as follows.

(1) If ϕ is an axiom of ZFC, then $^S\phi$ is an axiom of NST. In other words, all the axioms of ZFC are valid in the universe of standard sets.

(2) In the universe of all sets, all the axioms of ZFC are valid except the regularity axiom.

(3) Every standard set is internal. Every element of an internal set is internal.

(4) (transfer) For every n -ary formula $\phi(x_1, \dots, x_n)$ of ZFC, we have

$$\forall x_1 \in S \dots \forall x_n \in S [^S\phi(x_1, \dots, x_n) \leftrightarrow ^I\phi(x_1, \dots, x_n)].$$

(5) (saturation) For every set of size at most that of S , the scheme corresponding to Axiom 3 (κ -saturation) in Section B holds. Therefore the scheme corresponding to Axiom 2 (enlargement) holds also.

(6) (standardization) For every set a included in a standard set, there exists a standard set b such that we have $\forall x \in S [x \in a \leftrightarrow x \in b]$.

This completes the description of NST.

Theorem (Nelson-Hrbacek-Kawai): NST is a conservative extension of ZFC. That is, let ϕ be a sentence of ZFC. If $^S\phi$ can be proved in NST, then ϕ can be proved in ZFC.

Corollary: If ZFC is consistent, so is NST.

References

- [1] A. Robinson, Non-standard analysis, North-Holland, 1966; second edition, 1974.
- [2] M. Saito, Ultraproducts and non-standard analysis (in Japanese), Tokyo Tosho, 1976.
- [3] C. C. Chang and H. J. Keisler, Model theory, North-Holland, 1973.
- [4] K. D. Stroyan and W. A. J. Luxemburg, Introduction to the theory of infinitesimals, Academic Press, 1976.
- [5] H. J. Keisler, Elementary calculus, Prindle, Weber & Schmidt, 1976.
- [6] H. J. Keisler, Foundations of infinitesimal calculus, Prindle, Weber & Schmidt, 1977.
- [7] M. Davis, Applied nonstandard analysis, Wiley, 1976.
- [8] M. Saito, On the non-standard representation of linear mappings from a function space, Comm. Math. Univ. Sancti Pauli, 26 (1977), 165–185.
- [9] P. A. Loeb, Conversion from nonstandard to standard measure spaces and applications in probability theory, Trans. Amer. Math. Soc., 211 (1975), 113–122.
- [10] R. M. Anderson, A nonstandard representation for Brownian motion and Itô integration, Israel J. Math., 25 (1976), 15–46.
- [11] R. M. Anderson, Star-finite probability theory, Dissertation, Yale Univ., 1977.
- [12] H. J. Keisler, An infinitesimal approach to stochastic analysis, Mem. Amer. Math. Soc. 297 (1984).
- [13] E. Nelson, Internal set theory, Bull. Amer. Math. Soc., 83 (1977), 1165–1198.
- [14] K. Hrbacek, Axiomatic foundations for nonstandard analysis, Fund. Math., 98 (1978), 1–19.
- [15] K. Čuda, A nonstandard set theory, Comm. Math. Univ. Carolinae, 17 (1976), 647–663.
- [16] T. Kawai, Nonstandard analysis by axiomatic method, Southeast Asian Conference on Logic, C.-T. Chong and M. J. Wicks (eds.), Elsevier, 1983.
- [17] N. J. Cutland, Nonstandard measure theory and its applications, Bull. London Math. Soc., 15 (1983), 529–589.

294 (II.9) Numbers

A. General Remarks

From counting, a primitive mental activity, came the **natural numbers** (\mathbf{N}) (\rightarrow Section B), which serve to denote the number of items or the order in which these items are arranged. We can extend this concept to define, step by step, the $^*\text{integers}$ (\mathbf{Z}), $^*\text{rational numbers}$ (\mathbf{Q}), $^*\text{real numbers}$ (\mathbf{R}), and $^*\text{complex numbers}$ (\mathbf{C}) (\rightarrow 355 Real Numbers; 74 Complex Numbers).

The extensions up to the rationals are carried out to attain a domain within which the operations of addition, subtraction, multiplication, and division, namely, the **four arithmetic operations** (or **rational operations**) can be performed indefinitely, with division by 0 the only exception. To develop the theory of natural numbers there are two well-known methods, one being G. Peano's system of axioms [4], which will be stated below, and the other being R. Dedekind's set-theoretic treatment [5]. The domain of rational numbers can be extended, taking continuity into consideration, to that of real numbers by several methods of which the best known are those of the Dedekind cut (1872) [8], which will be described below, and of G. Cantor's

fundamental sequences (1872) [9]. There is also a way based on infinite series, given by K. Weierstrass in his lectures (1859–1860). Though in the domain of real numbers (1) the four arithmetic operations can be performed indefinitely and (2) an order relation for magnitude is defined, the equation $x^2 + 1 = 0$ has no root. Introducing numbers expressible as $a + ib$ ($i = \sqrt{-1}$), it is possible to solve every equation of the second order. Such numbers, once called imaginary, have been used since the days of G. Cardano [6] in the 16th century. L. Euler also made good use of complex numbers as convenient tools in many calculations and obtained among other things his formula $\exp i\theta = \cos \theta + i \sin \theta$. Indeed, the notation $i = \sqrt{-1}$ was used for the first time by Euler (1777). Furthermore, C. F. Gauss, giving imaginary numbers the name *complex numbers*, showed that any algebraic equation with numerical coefficients always has roots in the domain of complex numbers. The discovery of the geometric representation of complex numbers by several mathematicians in the late 18th and early 19th centuries and their use in many applications have made complex numbers indispensable in mathematics.

Though there are extensions of complex numbers, such as [†]Hamilton's quaternions or [†]Cayley numbers, it is generally accepted that when we speak of a *number* we usually mean a complex number.

B. Natural Numbers

Peano, basing his system on a specific natural number, 1, and a function such that to each natural number x corresponds a natural number $x + 1$ (hereafter denoted by x' and called the **successor** of x), formulated the fundamental properties of the set \mathbf{N} of natural numbers in the following five axioms, called the **Peano postulates**: (1) $1 \in \mathbf{N}$; (2) if $x \in \mathbf{N}$, then $x' \in \mathbf{N}$; (3) if $x \in \mathbf{N}$, then $x' \neq 1$; (4) if $x' = y'$ ($x, y \in \mathbf{N}$), then $x = y$; and (5) if a set M satisfies the two conditions: $1 \in M$, and $x \in M$ implies $x' \in M$, then $\mathbf{N} \subset M$. Since these postulates determine \mathbf{N} uniquely up to isomorphism, they can be regarded as a definition of \mathbf{N} . Elements of \mathbf{N} are called **natural numbers**.

Owing to Peano's fifth postulate, regarding a certain property $P(n)$ for natural numbers n we can deduce that $P(n)$ is true for every n if we prove both of the following conditions: (i) $P(1)$ is true; and (ii) for any natural number k , if $P(k)$ is true, then $P(k + 1)$ is true. Such reasoning is called **mathematical induction** (or **complete induction**). Accordingly, Peano's fifth postulate is called the **axiom of mathematical induction**. **Double mathematical induction**, a

generalization of mathematical induction, is as follows: to show that the property $P(m, n)$ is true for every pair of natural numbers m and n , we have only to show that (iii) $P(m, 1)$ and $P(1, n)$ are true for every m and every n ; and (iv) for every pair of natural numbers k and l , if $P(k + 1, l)$ and $P(k, l + 1)$ are true, then $P(k + 1, l + 1)$ is true. This axiom can be formulated in several other ways and can be generalized further to n -tuple mathematical inductions ($n = 2, 3, 4, \dots$), generically called **multiple mathematical inductions**.

Assume for a set M that a mapping f from the [†]Cartesian product $\mathbf{N} \times M$ into M is given. Then a mapping φ from \mathbf{N} into M such that (v) $\varphi(1) = a$; and (vi) $\varphi(x') = f(x, \varphi(x))$ ($x \in \mathbf{N}$) exists and is unique. Defining φ by (v) and (vi) is called the **definition of φ by mathematical induction**.

In particular, given a natural number a , the mapping $\varphi: \mathbf{N} \rightarrow \mathbf{N}$ defined by (vii) $\varphi(1) = a'$; and (viii) $\varphi(x') = \varphi(x)'$ is called **addition by a** . We shall write $\varphi(b) = a + b$, whence $x' = x + 1$. Addition thus defined obeys the following laws: $a + b = b + a$ (**commutative law**); $(a + b) + c = a + (b + c)$ (**associative law**). Peano's postulates are thus equivalent to the following: (1') $1 \in \mathbf{N}$; (2') for each pair $a, b \in \mathbf{N}$, $a + b \in \mathbf{N}$ is defined so that addition obeys the commutative and associative laws; (3') for each pair of natural numbers a and b , one and only one of the following three relations holds: $a = b + c$ ($c \in \mathbf{N}$); $a = b$; $a + c = b$ ($c \in \mathbf{N}$); (4') the same as (5) (mathematical induction). From (1')–(4') follows the **cancellation law**: $a + c = b + c \Leftrightarrow a = b$. Define $a \geq b$ if and only if $a = b$ or $a = b + c$ ($a, b, c \in \mathbf{N}$). Then, from (3'), \mathbf{N} becomes a [†]totally ordered set and $a \geq b \Leftrightarrow a + c \geq b + c$.

For each $a \in \mathbf{N}$, the mapping $\varphi: \mathbf{N} \rightarrow \mathbf{N}$ defined by $\varphi(1) = a$ and $\varphi(x') = \varphi(x) + a$ is called **multiplication by a** , and we write $\varphi(b) = ab$ (or $a \cdot b$). Multiplication obeys the following laws: $ab = ba$ (**commutative law**); $(ab)c = a(bc)$ (**associative law**); $a(b + c) = ab + ac$, $(a + b)c = ac + bc$ (**distributive laws**); and $ac = bc \Leftrightarrow a = b$ (**cancellation law**). The statement $a \cdot 1 = 1 \cdot a = a$ also holds.

Natural numbers, which have been introduced thus far as [†]ordinal numbers, also have the properties of [†]cardinal numbers. Denoting $\{1, 2, \dots, n\} = M_n$, we have $\overline{M_n} = \overline{M_m} \Leftrightarrow m = n$, $\overline{M_m} + \overline{M_n} = \overline{M_{m+n}}$, $\overline{M_m} \times \overline{M_n} = \overline{M_{mn}}$ (\rightarrow 49 Cardinal Numbers).

C. Integers

Introducing new numbers which are not in $\mathbf{N} = \{1, 2, \dots\}$, represented by the notations $0, -1, -2, \dots, -n, \dots$, we write $\mathbf{Z} = \{\dots,$

$-n, \dots, -2, -1, 0, 1, 2, \dots, n, \dots\}$. An element of \mathbf{Z} is called an **integer** (or **rational integer**).

Algebraically we can construct \mathbf{Z} from \mathbf{N} as follows: Let the set of all \dagger ordered pairs (k, l) of natural numbers k, l be $M = \mathbf{N} \times \mathbf{N}$, and define in M an \dagger equivalence relation $(k, l) \sim (m, n)$ by $k + n = m + l$. Let the equivalence class of (k, l) be $K(k, l)$, and construct the \dagger quotient space, $M/\sim = M^*$. Then the mapping $\varphi: \mathbf{Z} \rightarrow M^*$ defined by $\varphi(n) = K(k + n, k)$, $\varphi(0) = K(k, k)$, $\varphi(-n) = K(k, k + n)$ is bijective. Setting $K(k, l) + K(m, n) = K(k + m, l + n)$, addition can be defined in M^* (and accordingly in \mathbf{Z}) which is an extension of that in \mathbf{N} . Since $K(k, l) - K(m, n) = K(k + n, l + m)$, subtraction can be defined in \mathbf{Z} . This makes \mathbf{Z} an \dagger Abelian group with respect to addition. An order relation in \mathbf{Z} is defined by $K(k, l) \geq K(m, n) \Leftrightarrow k + n \geq m + l$, which makes \mathbf{Z} a totally ordered set. This order relation is an extension of that in \mathbf{N} . In particular, $\mathbf{N} = \{a \in \mathbf{Z} | a > 0\}$. Furthermore, setting $K(k, l) \times K(m, n) = K(km + ln, kn + lm)$, multiplication in \mathbf{Z} can be defined. It is an extension of that in \mathbf{N} and obeys commutative, associative, and distributive laws. Also, for each $a, b \in \mathbf{Z}$, we have $ab = 0 \Leftrightarrow (a = 0 \text{ or } b = 0)$. Thus, \mathbf{Z} becomes an \dagger integral domain.

D. Rational Numbers

Let P be the set of all ordered pairs (a, b) of integers a, b with $b \neq 0$, and define in P an equivalence relation $(a, b) \sim (c, d)$ by $ad = bc$. Each equivalence class determined by this relation is called a **rational number**. Denoting by $L(a, b)$ the equivalence class to which (a, b) belongs, we can define the sum $x + y$, the difference $x - y$, the product xy , and the quotient x/y of rational numbers $x = L(a, b)$, $y = L(c, d)$, as in the cases of addition and multiplication of integers, in the following way: $x + y = L(ad + bc, bd)$, $x - y = L(ad - bc, bd)$, $xy = L(ac, bd)$, $x/y = L(ad, bc)$, where the quotient is defined only when $c \neq 0$. Thus the set \mathbf{Q} of all rational numbers becomes a \dagger field.

For the same reason that we have identified integers of a special type with natural numbers, we now identify a rational number expressible as $L(a, 1)$ with an integer a . Henceforth, any rational number $L(a, b)$ ($b \neq 0$) can be expressed in the form of a quotient a/b ($b \neq 0$) of integers a and b .

From $L(a, b) = L(ca, cb)$ ($c \neq 0$), it is always possible to assume $b > 0$ in the representation $L(a, b)$ of a rational number x . For any two rational numbers $x = L(a, b)$, $y = L(c, d)$ with $b > 0$, $d > 0$, define an ordering in \mathbf{Q} by $x \geq y \Leftrightarrow ad \geq bc$, which is an extension of the ordering of integers. Thus (i) \mathbf{Q} becomes totally

294 E Numbers

ordered, and we have (ii) $x \geq y \Rightarrow x + z \geq y + z$ and (iii) $x \geq y$ and $z \geq 0 \Rightarrow xz \geq yz$. The rational x is called **positive** if $x > 0$ and **negative** if $x < 0$.

E. Real Numbers

Two typical methods of constructing real numbers from rational numbers are those of Dedekind and of Cantor.

Dedekind's Theory of Real Numbers. We call a pair (A_1, A_2) of subsets A_1, A_2 of the set \mathbf{Q} of all rational numbers a **cut** of \mathbf{Q} if they satisfy the following conditions: (i) $A_1 \neq \emptyset$, $A_2 \neq \emptyset$; (ii) $\mathbf{Q} = A_1 \cup A_2$; (iii) $a_1 \in A_1, a_2 \in A_2 \Rightarrow a_1 < a_2$. Then the following three cases are distinguished: there is (i) a maximum in A_1 with no minimum in A_2 ; (ii) no maximum in A_1 with a minimum in A_2 ; or (iii) no maximum in A_1 with no minimum in A_2 . A cut with either condition (i) or (iii) is called a **real number** (in the sense of Dedekind); condition (ii) can be converted to (i). The set of all real numbers is denoted hereafter by \mathbf{R} and each real number by α or β or A real number with property (i) is called a **rational real number**, and a real number with property (iii) an **irrational real number**. Any rational real number is uniquely determined by the maximum a^* of A_1 , and the mapping: $a \rightarrow a^* = (A_1, A_2)$ from the set \mathbf{Q} of rational numbers onto the set \mathbf{Q}^* of rational real numbers is bijective.

I. For real numbers $\alpha = (A_1, A_2)$ and $\beta = (B_1, B_2)$ we define $\alpha \leq \beta$ if and only if $A_1 \subset B_1$. By this ordering \leq , \mathbf{R} becomes a totally ordered set.

II. For real numbers $\alpha = (A_1, A_2)$ and $\beta = (B_1, B_2)$, put $C_2 = \{a + b | a \in A_2, b \in B_2\}$ and $C_1 = \mathbf{R} - C_2$; then $(C_1, C_2) = \gamma$ is a real number. Define the sum $\alpha + \beta$ by setting $\alpha + \beta = \gamma$. Addition thus defined obeys commutative and associative laws, and \mathbf{R} becomes an \dagger Abelian group with 0^* as its zero element. Furthermore, for real numbers $\alpha = (A_1, A_2)$ and $\beta = (B_1, B_2)$ with $0^* \leq \alpha$, $0^* \leq \beta$, put $D_2 = \{ab | a \in A_2, b \in B_2\}$, $D_1 = \mathbf{R} - D_2$; then $(D_1, D_2) = \delta$ becomes a real number. Define the product $\alpha\beta$ by setting $\alpha\beta = \delta$. According as $0^* > \alpha$, $0^* \leq \beta$; $0^* \leq \alpha$, $0^* > \beta$; or $0^* > \alpha$, $0^* > \beta$, define $\alpha\beta = -((- \alpha)\beta)$; $\alpha\beta = -(\alpha(-\beta))$; and $\alpha\beta = (-\alpha)(-\beta)$, respectively. Multiplication thus defined obeys the commutative, associative, and distributive laws, and \mathbf{R} becomes a field with 1^* as its unity element.

III. For ordering and arithmetic operations, we have (1) $\alpha \geq \beta \Rightarrow \alpha + \gamma \geq \beta + \gamma$; and (2) $\alpha \geq \beta$, $\gamma \geq 0^* \Rightarrow \alpha\gamma \geq \beta\gamma$.

By letting each rational number a corre-

spond to a rationally real number a^* , we can set up a bijection between the set \mathbf{Q} of all rational numbers and \mathbf{Q}^* . Furthermore, in this correspondence, the sum, product, 0, and 1 of \mathbf{Q} are mapped to the sum, product, zero element, and unity element of \mathbf{Q}^* , respectively; in addition, the ordering is preserved. Thus \mathbf{Q} and \mathbf{Q}^* are isomorphic with respect to both arithmetic operations and ordering. We shall hereafter identify the element a^* with a . Accordingly we call rational real numbers simply **rational numbers** and similarly irrational real numbers simply **irrational numbers**.

IV. Similarly as for \mathbf{Q} , we can define a cut (A_1, A_2) of the set of all real numbers \mathbf{R} (precisely, by the conditions $A_1 \neq \emptyset$, $A_2 \neq \emptyset$; $\mathbf{R} = A_1 \cup A_2$; $d_1 \in A_1, d_2 \in A_2 \Rightarrow d_1 < d_2$). For each cut (A_1, A_2) of \mathbf{R} , either there is a maximum in A_1 and no minimum in A_2 or no maximum in A_1 and a minimum in A_2 (Dedekind). This property is called the **connectedness** (or **continuity**) of real numbers.

By using definitions (I)–(IV), any real number α can be represented (i) as the †least upper bound of some set A of rational numbers: $\alpha = \sup A$; and also (ii) as the limit of a sequence $\{a_n\}$ of rational numbers: $\alpha = \lim a_n$.

Cantor's Theory of Real Numbers. A sequence $\{a_n\}$ of rational numbers is called a **fundamental sequence** (or **Cauchy sequence**) if it satisfies the following condition: for each positive rational number ε , there exists a natural number n_0 such that $-\varepsilon < a_n - a_m < \varepsilon$ for every a_m and a_n with $m > n_0$ and $n > n_0$. For two Cauchy sequences $\{a_n\}$ and $\{b_n\}$, we can write $\{a_n\} \sim \{b_n\}$ if $\{a_1, b_1, a_2, b_2, \dots, a_n, b_n, \dots\}$ is again a Cauchy sequence. The relation \sim is an equivalence relation. Let \mathbf{R}' be the set of all equivalence classes obtained by classifying, with respect to \sim , the set of all Cauchy sequences, and call an element of \mathbf{R}' (an equivalence class) a **real number** (in the sense of Cantor). We shall write $[\{a_n\}]$ hereafter for the equivalence class of $\{a_n\}$. In particular, $\{a_n\}$ with $a_n = a$ ($n = 1, 2, \dots$) being a Cauchy sequence, we denote a real number $[\{a_n\}]$ by a^{**} . An element of \mathbf{R}' of this type is called a **rational real number** (in the sense of Cantor), while one not of this type is called an **irrational real number** (in the sense of Cantor).

For each pair of real numbers $\alpha = [\{a_n\}]$ and $\beta = [\{b_n\}]$, their sum and product are uniquely defined by $\alpha + \beta = [\{a_n + b_n\}]$ and $\alpha\beta = [\{a_nb_n\}]$, where both $\{a_n + b_n\}$ and $\{a_nb_n\}$ are shown to be Cauchy sequences. Further, for α and β , define $\alpha < \beta$ if $a_n < b_n$ hold for all n larger than some number n_0 . As regards arithmetic operations and ordering, \mathbf{R}' has the properties (I)–(IV) of Dedekind's theory. Further, any Cauchy sequence of which the terms

all are real numbers has a limit in \mathbf{R}' (**completeness of real numbers**).

The sets of real numbers obtained by the above two methods are isomorphic with respect to arithmetic operations and ordering (\rightarrow 355 Real Numbers).

F. Complex Numbers

To construct complex numbers from real ones, several methods have been devised, among which the one stated below is due to W. R. Hamilton.

An ordered pair (a, b) of real numbers a and b will be called a **complex number**. Arithmetic operations among complex numbers are defined as follows: $(a, b) + (c, d) = (a + c, b + d)$, $(a, b) - (c, d) = (a - c, b - d)$, $(a, b) \cdot (c, d) = (ac - bd, bc + ad)$, and

$$\frac{(a, b)}{(c, d)} = \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2} \right), \quad (c, d) \neq (0, 0).$$

According to these definitions, the set \mathbf{C} of all complex numbers becomes a †field with $(0, 0)$ and $(1, 0)$ as its †zero element and †unity element, respectively. $\mathbf{R}^* = \{(a, 0) | a \in \mathbf{R}\}$ is a †subfield of \mathbf{C} , and the mapping $\varphi: \mathbf{R} \rightarrow \mathbf{R}^*$ defined by $\varphi(a) = (a, 0)$ proves to be a field †isomorphism. Thus, identifying the element $(a, 0)$ of \mathbf{R}^* with the element a of \mathbf{R} , \mathbf{C} can be regarded as an †overfield of \mathbf{R} ($\mathbf{R} \subset \mathbf{C}$). Hence, the zero element of \mathbf{C} is the real number 0 and the unity element is the real number 1. The complex number $(0, 1)$ is called the **imaginary unit** and denoted by i . Thus we can write $a + bi$ as usual for a complex number (a, b) (\rightarrow 74 Complex Numbers).

References

- [1] O. Perron, *Irrationalzahlen*, de Gruyter, second edition, 1929 (Chelsea, 1948).
- [2] E. G. H. Landau, *Grundlagen der Analysis*, Akademische Verlag, 1934; English translation, *Foundations of analysis*, Chelsea, 1960.
- [3] N. Bourbaki, *Eléments de mathématique*, I. *Théorie des ensembles*, ch. 3, *Actualités Sci. Ind.*, 1243b, second edition, 1967; III. *Topologie générale*, ch. 4, 1143c, third edition, 1960; ch. 8, 1235b, third edition, 1963; English translation, *Theory of sets*, Addison-Wesley, 1968; *General topology*, pt. I, Addison-Wesley, 1966.
- [4] G. Peano, *Arithmetices principia nova methodo exposita*, Bocca, 1889 (*Opere scelte* II, Cremona, 1958).
- [5] R. Dedekind, *Was sind und was sollen die Zahlen?* Vieweg, 1887 (*Gesammelte mathematische Werke* 3, Vieweg, 1932); English translation, *Essays on the theory of numbers*, Open Court, 1901.

- [6] G. Cardano, *Artis magnae sive de regulis algebraicis*, Nuremberg, 1545.
 [7] W. R. Hamilton, *Lectures on quaternions*, Hodges and Smith, 1853.
 [8] R. Dedekind, *Stetigkeit und irrationale Zahlen*, Vieweg, 1872 (Gesammelte mathematische Werke 3, Vieweg, 1932); English translation, *Essays on the theory of numbers*, Open Court, 1901.
 [9] G. Cantor, *Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen*, *Math. Ann.*, 5 (1872), 123–132 (Gesammelte Abhandlungen, Springer, 1932).

295 (V.4) Number-Theoretic Functions

A. Recurrent Sequences

A (complex-valued) function that has the set of nonnegative integers (or the set \mathbb{N} of natural numbers) as its domain is called a **number-theoretic** (or **arithmetic**) **function**. Thus it can be regarded as a sequence of numbers. We first consider recurrent sequences. Let $f(x_0, x_1, \dots, x_{r-1})$ be a complex-valued function of r variables. Put $u_0 = a_0$, $u_1 = a_1, \dots, u_{r-1} = a_{r-1}$, and successively define $u_{i+r} = f(u_i, u_{i+1}, \dots, u_{i+r-1})$ ($i = 0, 1, 2, \dots$). The sequence $\{u_n\}$ thus defined is called a **recurrent sequence of order r** determined by the initial values a_0, a_1, \dots, a_{r-1} and the function f . In particular, when the defining function f is given by $\sum_{i=0}^{r-1} b_i x_i$, $\{u_n\}$ is called a **linear recurrent sequence**. The **Fibonacci sequence** is a special linear recurrent sequence with initial values a_0, a_1 and the defining function $x_0 + x_1$. Let $\alpha = (1 + \sqrt{5})/2$, $\beta = (1 - \sqrt{5})/2$ be two roots of $1 + x = x^2$, and let c_1 and c_2 be determined by $c_1 + c_2 = a_0$ and $c_1\alpha + c_2\beta = a_1$. Then the Fibonacci sequence with the initial values a_0, a_1 is given by putting $u_n = c_1\alpha^n + c_2\beta^n$ ($n = 0, 1, 2, \dots$). If we put $a_0 = 1$, $a_1 = 1$, then we obtain **Binet's formula**:

$$u_n = \frac{1}{\sqrt{5}} \left\{ \left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1} \right\}.$$

B. Multiplicative Functions

A number-theoretic function f with domain \mathbb{N} is said to be **multiplicative** if $f(1) = 1$, $f(mn) = f(m)f(n)$ for $(m, n) = 1$, and to be **completely multiplicative** if $f(1) = 1$, $f(mn) = f(m)f(n)$ for any m and n ($\in \mathbb{N}$). Similarly, f is said to be **additive** if $f(1) = 0$, $f(mn) = f(m) + f(n)$ for $(m, n) = 1$, and to be **completely additive** if $f(1) = 0$

and $f(mn) = f(m) + f(n)$ for any m and n ($\in \mathbb{N}$). If f is multiplicative and $\sum_{n=1}^{\infty} f(n)$ is absolutely convergent, then $\sum_{v=0}^{\infty} f(p^v)$ is absolutely convergent for any prime p . Moreover, we have

$$\sum_{n=1}^{\infty} f(n) = \prod_p \left(\sum_{v=0}^{\infty} f(p^v) \right),$$

where the infinite product on the right-hand side is also absolutely convergent. Furthermore, if $f(n)$ is completely multiplicative, then we have

$$\sum_{n=1}^{\infty} f(n) = \prod_p (1 - f(p))^{-1}.$$

In particular, for $f(n) = n^{-s}$, which is completely multiplicative (with $s = \sigma + it$ a complex variable), we get the \dagger Euler product formula for the \dagger Riemann zeta function:

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_p (1 - p^{-s})^{-1}$$

for $\sigma > 1$.

C. Convolutions

If f and g are number-theoretic functions, the **convolution $f * g$** is defined by

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d),$$

with the summation carried over all divisors d of n . For any number-theoretic functions f, g, h , we have $f * g = g * f$ and $(f * g) * h = f * (g * h)$. If f and g are multiplicative number-theoretic functions, $f * g$ is again a multiplicative number-theoretic function.

The **Möbius function** $\mu(n)$ is defined as follows: $\mu(1) = 1$, $\mu(n) = 0$ if n is divisible by the square of a prime, and $\mu(n) = (-1)^r$ if n is the product of r distinct primes. It can easily be proved that μ is multiplicative and that

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & (n = 1), \\ 0 & (n > 1). \end{cases}$$

Let e and ρ be functions defined by $e(1) = 1$, $e(n) = 0$ ($n > 1$), and $\rho(n) = 1$ for every n . Then e is the identity element for the convolution $*$, and $\mu * \rho = \rho * \mu = e$. It follows that $f * \rho = g$ is equivalent to $f = g * \mu$, that is,

$$g(n) = \sum_{d|n} f(d)$$

implies that

$$f(n) = \sum_{d|n} \mu(d)g(n/d),$$

and vice versa. We call the latter the **Möbius inversion formula**. Similarly, for complex-valued functions F, G defined on $[1, +\infty)$,

$$G(x) = \sum_{n \leq x} F(x/n)$$

Number-Theoretic Functions

is equivalent to

$$F(x) = \sum_{n \leq x} \mu(n) G(x/n).$$

Let $\varphi(n)$ be the number of integers m not greater than n and such that $(m, n) = 1$. The function $\varphi(n)$, called the **Euler function**, is multiplicative, and we have

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right) = n \sum_{d|n} \frac{\mu(d)}{d},$$

and $\varphi(p^n) = p^n - p^{n-1}$ for every prime p . Let v be the function defined by $v(n) = n$ for every n . Then we have $\mu * v = \varphi$ and $v = \rho * \varphi$, and hence $\sum_{d|n} \varphi(d) = n$.

The **generalized divisor function** $d_k(n)$ is the number of ways of expressing n as a product of k factors. Thus we have

$$d_k(n) = \sum_{n_1 n_2 \dots n_k = n} 1 = \sum_{m|n} d_{k-1}(m),$$

and $d_k = \rho * \dots * \rho$ (k factors). Therefore, d_k is multiplicative. For simplicity, we write $d(n)$ instead of $d_2(n)$. Then $d(n)$ is the number of divisors of n , and we call it the **divisor function**. For example, $d(12) = 6$. We denote by $\sigma_\alpha(n)$ the sum of α th powers of divisors of n . If $n = \prod_{p|n} p^l$, then we have

$$\sigma_\alpha(n) = \sum_{d|n} d^\alpha = \prod_{p|n} \frac{p^{\alpha(l+1)} - 1}{p^\alpha - 1},$$

and σ_α is also multiplicative. We write $\sigma_1(n) = \sigma(n)$. A number n satisfying $\sigma(n) = 2n$ is said to be a † perfect number. Let $n = \prod_{p|n} p^{l_p}$. The functions $\omega(n) = \sum_{p|n} 1$ (the number of distinct prime factors of n) and $\Omega(n) = \sum_{p|n} l_p$ (the number of prime factors of n) are often used, the former being additive and the latter completely additive.

D. Residue Characters and Gaussian Sums

Let k be a positive integer and χ be a completely multiplicative function such that $\chi(n) = 0$ for $(n, k) > 1$ and $\chi(n_1) = \chi(n_2)$ for $n_1 \equiv n_2 \pmod{k}$. The function χ is called a **Dirichlet character** (or **residue character**) with modulus k (or modulo k). There exist $\varphi(k)$ distinct characters modulo k for a given k . The character satisfying $\chi(n) = 1$ for every n coprime to k is called the **principal character** modulo k , and we denote it by χ_0 . It is easily proved that

$$\sum_{n=0}^{k-1} \chi(n) = \begin{cases} \varphi(k) & (\chi = \chi_0), \\ 0 & (\chi \neq \chi_0), \end{cases}$$

$$\sum_{\chi} \chi(n) = \begin{cases} \varphi(k) & (n \equiv 1 \pmod{k}), \\ 0 & (n \not\equiv 1 \pmod{k}). \end{cases}$$

Suppose that $k = k_1 k_2 \dots k_r$ ($(k_i, k_j) = 1, i \neq j$). Then there exist r characters χ_i modulo k_i permitting the unique decomposition $\chi = \chi_1 \chi_2 \dots \chi_r$.

Let χ be a character modulo k and $\rho = \exp(2\pi i/k)$. The **Gaussian sum** modulo k is defined by

$$G(a, \chi) = \sum_{n \pmod{k}} \chi(n) \rho^{an},$$

where n runs over a complete system modulo m (\rightarrow 297 Number Theory, Elementary, G). Hence if $a \equiv b \pmod{k}$, then $G(a, \chi) = G(b, \chi)$. Suppose that $k = k_1 k_2 \dots k_r$ ($(k_i, k_j) = 1, i \neq j$) and $a_i \equiv a \pmod{k_i}$. Then we have

$$G(a, \chi) = \varepsilon G(a_1, \chi_1) G(a_2, \chi_2) \dots G(a_r, \chi_r),$$

$$\varepsilon = \chi_1(k/k_1) \chi_2(k/k_2) \dots \chi_r(k/k_r),$$

where $\chi = \chi_1 \chi_2 \dots \chi_r$ is the decomposition of χ mentioned above.

Let k_0 be a divisor of k . If $\chi(n) = 1$ whenever $(n, k) = 1$ and $n \equiv 1 \pmod{k_0}$, then we say that χ is defined mod k_0 . The least positive integer $f = f(\chi)$ modulo to which χ is defined is called the **conductor** of χ . If the conductor of χ is k itself, then χ is said to be a **primitive character** modulo k . When χ has the decomposition $\chi = \chi_1 \dots \chi_r$, the conductor of χ is also decomposed as $f(\chi) = f(\chi_1) f(\chi_2) \dots f(\chi_r)$. If D is a square-free integer, then the † discriminant d of the quadratic number field $\mathbf{Q}(\sqrt{D})$ (where \mathbf{Q} is the rational number field) is either D ($D \equiv 1 \pmod{4}$) or $4D$ ($D \equiv 2, 3 \pmod{4}$). The integers d so represented are called the **fundamental discriminants**. The † Kronecker symbol (d/n) (\rightarrow 347 Quadratic Fields) is defined only for such fundamental discriminants. Z. Suetuna and A. Z. Walfisz (1936) proved that if $\chi(n)$ is a real primitive character modulo k , then we necessarily have one of the following cases: (i) $k = p_1 p_2 \dots p_r$; (ii) $k = 4p_1 p_2 \dots p_r$; or (iii) $k = 8p_1 p_2 \dots p_r$ (with the p_i distinct odd primes). In case (i),

$$\chi(n) = \begin{cases} (k/n) & (k \equiv 1 \pmod{4}), \\ (-k/n) & (k \equiv -1 \pmod{4}); \end{cases}$$

in case (ii),

$$\chi(n) = \begin{cases} (-k/n) & (k/4 \equiv 1 \pmod{4}), \\ (k/n) & (k/4 \equiv -1 \pmod{4}); \end{cases}$$

and in case (iii), $\chi(n)$ is either (k/n) or $(-k/n)$.

If χ is a primitive character modulo k , then $G(a, \chi) = \bar{\chi}(a) G(1, \chi)$, $G(1, \chi) \bar{G}(1, \chi) = k$. In particular, if χ is a real primitive character, then

$$G(1, \chi) = \begin{cases} \sqrt{k} & (\chi(-1) = 1), \\ i\sqrt{k} & (\chi(-1) = -1). \end{cases}$$

Sometimes $S(a, k) = \sum_{n \pmod{k}} \rho^{an^2}$ is called the Gaussian sum, where $\rho = \exp(2\pi i/k)$. If $(a, k_i) = 1$ ($i = 1, 2$) and $(k_1, k_2) = 1$, then $S(a, k_1 k_2) = S(ak_2, k_1) S(ak_1, k_2)$. If $(a, 2) = 1$, then

$$S(a, 2^r) = \begin{cases} 0 & (r = 1), \\ (1 + i^a) 2^{r/2} & (r \text{ even}), \\ 2^{\pi i a/4} \cdot 2^{(r+1)/2} & (r \text{ odd } > 1). \end{cases}$$

Let p be an odd prime and $(a, p) = 1$. Then we have

$$S(a, p^r) = \begin{cases} (a/p)S(1, p) & (r=1), \\ p^{r/2} & (r \text{ even}), \\ p^{(r-1)/2}S(a, p) & (r \text{ odd} > 1), \end{cases}$$

where (a/p) is the \dagger Legendre symbol (\rightarrow 297 Number Theory, Elementary, H). The well-known **Gauss formula** is stated as follows:

$$S(1, p) = \begin{cases} \sqrt{p} & (p \equiv 1 \pmod{4}), \\ i\sqrt{p} & (p \equiv 3 \pmod{4}). \end{cases}$$

Various proofs of this formula have been given by many authors [5].

Let

$$G(a, \chi_0) = \sum'_{h \pmod{k}} \exp(2\pi i ah/k),$$

called the **Ramanujan sum**, be denoted by $c_k(a)$. The sum $\sum'_{h \pmod{k}}$ means that h runs through a reduced residue system modulo m (\rightarrow 297 Number Theory, Elementary, G). It follows that $c_k(1) = \mu(k)$ and

$$c_k(a) = \sum_{d|(k,a)} \mu\left(\frac{k}{d}\right)d.$$

E. Analytic Methods

Let f be a number-theoretic function. One of the problems in analytic number theory is to estimate $\sum_{n \leq x} f(n)$ or to expand it into series. We assume that $f(x)$ and $g(x)$ are real-valued functions defined for $x \geq 1$ and $g(x)$ is of class C^1 . Then we have

$$\sum_{n \leq x} f(n)g(n) = F(x)g(x) - \int_1^x F(t)g'(t)dt,$$

where $F(x) = \sum_{n \leq x} f(n)$. If we take $f(x) = 1$, $g(x) = \log x$ or $1/x$ in this formula, then we obtain

$$\sum_{n \leq x} \log n = x \log x - x + O(\log x)$$

or

$$\sum_{n \leq x} n^{-1} = \log x + C + O(x^{-1})$$

(where C is the \dagger Euler constant), respectively.

We now construct from the function f the Dirichlet series

$$F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$$

or the power series

$$P(x) = \sum_{n=1}^{\infty} f(n)x^n.$$

There are called the **generating functions** of f . The consideration of these functions makes possible the application of function-theoretic

methods to the problem in this section. Sometimes more complicated functions are used.

We mainly consider the generating function represented by the Dirichlet series. The generating functions of ρ , μ , d , χ (\rightarrow Sections C, D) are $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$, $\sum_{n=1}^{\infty} \mu(n)n^{-s}$, $\sum_{n=1}^{\infty} d(n)n^{-s}$, $L(s, \chi) = \sum_{n=1}^{\infty} \chi(n)n^{-s}$, respectively. The \dagger abscissa of absolute convergence of each of these Dirichlet series is 1. The function $L(s, \chi)$ is called the \dagger Dirichlet L -function (\rightarrow 123 Distribution of Prime Numbers; 450 Zeta Functions). When $k=1$ and $\chi = \chi_0$, $L(s, \chi)$ is precisely $\zeta(s)$.

Let $F(s)$, $G(s)$, and $H(s)$ be the generating functions of f , g , and $h = f * g$. Let f and g be multiplicative, so that h is also multiplicative. Moreover, if $F(s)$ and $G(s)$ are absolutely convergent for $\sigma \geq \sigma_0$, then $H(s)$ is also absolutely convergent for $\sigma \geq \sigma_0$, and $F(s)G(s) = H(s)$ for $\sigma \geq \sigma_0$. It follows from this that $\sum_{n=1}^{\infty} \mu(n)n^{-s} = \zeta^{-1}(s)$, $\sum_{n=1}^{\infty} \mu(n)\chi(n)n^{-s} = L(s, \chi)^{-1}$, $\sum_{n=1}^{\infty} d_k(n)n^{-s} = \zeta(s)^k$, and $\sum_{n=1}^{\infty} d^2(n)n^{-s} = \zeta^4(s)/\zeta(2s)$. The last result was given by S. Ramanujan. More generally, $\sum_{n=1}^{\infty} d^r(n)n^{-s} = \zeta^{2r}(s)\varphi(s)$, where $\varphi(s)$ is an analytic function for $\sigma > 1/2$. By utilizing analytic methods, we can deduce from this that $\sum_{n \leq x} d^r(n) \sim x(c_1(\log x)^{2^r-1} + \dots + c_{2^r})$ (B. M. Wilson, 1923) (\rightarrow 123 Distribution of Prime Numbers; 242 Lattice-Point Problems).

There are many formulas known as the **Euler summation formula**. Among them the following one is convenient to use (E. Landau, L. J. Mordell, H. Davenport): Let $f_0(x) = x - [x] - 1/2$ (when x is not an integer), $= 0$ (when x is an integer). We successively construct continuous functions $f_1(x)$, $f_2(x)$, ... of period 1 such that $f'_r(x) = f_{r-1}(x)$ ($r \geq 1$) if x is not an integer, and $\int_0^1 f_r(x)dx = 0$. If $F(x)$ is of class C^h on $[a, b]$, then using these auxiliary functions, we have

$$\begin{aligned} \sum'_{a \leq m \leq b} F(m) &= \int_a^b F(x)dx \\ &+ \sum_{r=0}^{h-1} (-1)^r (f_r(a)F^{(r)}(a) \\ &\quad - f_r(b)F^{(r)}(b)) \\ &+ (-1)^{h-1} \int_a^b f_{h-1}(x)F^{(h)}(x)dx, \end{aligned}$$

where \sum' means that the term corresponding to $m=a$ or $m=b$ in the sum is to be replaced by $F(a)/2$ or $F(b)/2$ whenever a or b is an integer. Since the \dagger Fourier series of $f_0(x)$ is

$$-\sum_{n=1}^{\infty} \frac{\sin(2\pi nx)}{n\pi}$$

(which is convergent), $f_r(x)$ can be expressed by

$$-\sum_{n=-\infty}^{\infty} \frac{\exp(2\pi inx)}{(2\pi in)^{r+1}},$$

where Σ' means that the term with $n=0$ is omitted and this sum actually means

$$\sum_{n=1}^{\infty} \left(\frac{\exp(2\pi i n x)}{(2\pi i n)^{r+1}} + \frac{\exp(-2\pi i n x)}{(-2\pi i n)^{r+1}} \right).$$

Hence if $h=1$, we have

$$\sum_{a \leq n \leq b} F(n) = \int_a^b F(x) dx - \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \int_a^b F'(x) \sin(2\pi n x) dx.$$

For instance, if we put $h=1$, $a=1$, $b=N$, $F(x)=x^{-s}$ and let N tend to ∞ , then we have the formula

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \frac{1}{s-1} + \frac{1}{2} - s \int_1^{\infty} \frac{f_0(x)}{x^{s+1}} dx \quad \text{for } \sigma > 1.$$

Utilizing the following integration by parts, the integral on the right-hand side can be extended so as to become holomorphic in s in the whole complex plane:

$$\int_1^{\infty} \frac{f_0(x)}{x^{s+1}} dx = -f_1(1) + (s+1) \int_1^{\infty} \frac{f_1(x)}{x^{s+2}} dx = \dots$$

Probabilistic considerations are also used for the study of various number-theoretic functions. If $f(n)$ is $\omega(n)$ or $\Omega(n)$, then we have $\sum_{n \leq x} f(n) = x \log \log x + cx + o(x)$. Therefore the average order of $\omega(n)$ or $\Omega(n)$ is estimated as $\log \log n$. Let $A(x; \alpha, \beta)$ be the number of n satisfying $n \leq x$ and $\log \log n + \alpha \sqrt{\log \log n} < f(n) < \log \log n + \beta \sqrt{\log \log n}$. Then

$$\lim_{x \rightarrow \infty} \frac{A(x; \alpha, \beta)}{x} = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-u^2/2} du.$$

For $f(n)=\omega(n)$, the result was proved by P. Erdős and M. Kac (1940) by using the †central limit theorem and V. Brun's †sieve method. Further general formulas were obtained by M. Tanaka (1955). Further development and the present state of probabilistic number theory can be seen in [7, 8].

Finally, we mention the well-known estimation formulas. Let ε be an arbitrary positive number and n be a positive integer. Then $d(n) = O(n^{\varepsilon})$, where O (†Landau's symbol) depends on ε . We know that $\limsup_{n \rightarrow \infty} (\log d(n)) \cdot (\log \log n) / \log n = \log 2$ (S. Wigert, 1907), and $\liminf_{n \rightarrow \infty} \omega(n) \cdot (\log \log n) / n = e^{-c}$. The result $\omega(n) = O(\log n / \log \log n)$ is often used. Let χ be a primitive character modulo k , and let $S_m = \sum_{n=1}^m \chi(n)$. We can prove that $|S_m| < \sqrt{k} \log k$ for all m (I. Schur, 1918) and $S_m = \Omega(\sqrt{k} \log \log k)$ (R. Paley, 1932). This formula, with the symbol

Ω , means that there exist infinitely many k satisfying $|S_m| > c\sqrt{k} \log \log k$ by taking m and χ suitably, with c any positive constant. Let χ be a character modulo k , and χ^0 be a primitive character associated with χ . Then we have $L(s, \chi) = L(s, \chi^0) \prod_{p|k} (1 - \chi^0(p)p^{-s})$. It follows that

$$\sum_{m=1}^n \chi(m) = \sum_{d: d|k, d|f} \mu(d) \chi^0(d) \sum_{m \leq n/d} \chi^0(m)$$

if the conductor of χ^0 is f . A. G. Postnikov investigated (1956) the sum of characters. The least integer that is not a quadratic residue modulo p does not exceed $p^{1/2} \sqrt{e} (\log p)^2$, where p is a sufficiently large prime (I. M. Vinogradov, 1926). The least integer that is a †primitive root modulo p does not exceed $2^{m+1} \sqrt{p}$, where m is the number of distinct prime divisors of $p-1$, with p a prime (L. K. Hua, 1942). D. A. Burgess (1962) deals with the latter results. We conclude with Artin's conjecture: If w is a given square-free integer, then there exist infinitely many primes p such that w is a primitive root modulo p . In 1967, C. Hooley proved this conjecture subject to the assumption that the general †Riemann hypothesis holds for †Dedekind zeta functions over certain †Kummer extensions of \mathbb{Q} . He also obtained the asymptotic formula for the number of such primes not exceeding x . Recently the †sieve method has been widely applied to the various investigations in the theory of number theoretic functions (\rightarrow 123 Distribution of Prime Numbers E).

References

- [1] R. G. Ayoub, An introduction to the analytic theory of numbers, Amer. Math. Soc. Math. Surveys, 1963.
- [2] A. O. Gel'fond and Yu. V. Linnik, Elementary methods in analytic number theory, Rand McNally, 1965. (Original in Russian, 1962.)
- [3] G. H. Hardy and E. M. Wright, An introduction to the theory of numbers, Clarendon Press, fourth edition, 1965.
- [4] H. Hasse, Vorlesungen über Zahlentheorie, Springer, 1950.
- [5] E. G. H. Landau, Vorlesungen über Zahlentheorie I, Hirzel, 1927 (Chelsea, 1946).
- [6] W. J. Leveque, Topics in number theory, Addison-Wesley, 1956.
- [7] J. Kubilius, Probabilistic methods in the theory of numbers, Amer. Math. Soc. Transl. Math. Monographs, 1964. (Original in Russian, 1962.)
- [8] P. D. T. A. Elliott, Probabilistic number theory I, II, Springer, 1979.

296 (V.1) Number Theory

A. History

Simple and curious relations among integers were discovered and admired from antiquity. For example, we have the relation $3^2 + 4^2 = 5^2$, which has geometric meaning concerning right triangles. The Pythagoreans (\rightarrow 187 Greek Mathematics) sought similar relations. They were also interested in \dagger perfect numbers (numbers equal to the sum of their divisors, such as $28 = 1 + 2 + 4 + 7 + 14$). Modern arithmetic inherits from the Greeks the proof of the existence of an infinite number of primes, the \dagger Euclidean algorithm to obtain the greatest common divisor of two integers (both given in Euclid's *Elements*), and Eratosthenes' \dagger sieve for finding primes. In the 3rd century A.D., Diophantus of Alexandria discovered a method of solving indeterminate equations of degrees 1 and 2; this marked the origin of **Diophantine analysis**. The ancient Chinese also knew how to solve equations of the first degree in some special cases. Arithmetic also developed in India from an early period; its relation to Greek mathematics is not yet entirely clear. In the 12th century, the Indian mathematician Bhāskara knew how to solve \dagger Pell equations using a method much like Lagrange's.

Interest in integers was reborn in Europe in the 17th century. During this period, Bachet de Méziriac (1581–1638) rediscovered the solution of the \dagger Diophantine equation of the first degree and published it in his famous book on mathematical recreations [1]. Primes of the form $2^p - 1$, closely related to perfect numbers and called \dagger Mersenne numbers, attracted considerable interest. \dagger Fermat, sometimes called the father of number theory, announced numerous results without giving proofs; the most famous among them is the so-called Fermat's last theorem (\rightarrow 145 Fermat's Problem). Another famous conjecture of his is that every integer is expressible as the sum of at most n **n -gonal numbers**, i.e., numbers of the form $k(k-1)n/2 - k(k-2)$, $k \in \mathbb{N}$. This was proved by Gauss (for the case $n=3$), Jacobi ($n=4$), and Cauchy (for the general case).

In the 18th century, remarkable progress was made by Euler and Lagrange. The second part of Euler's *Algebra* [2] contains rich results of miscellaneous sorts in the field. Lagrange developed the theory of \dagger continued fractions and applied it to arithmetic. Toward the end of the 18th century, Legendre compiled his comprehensive book [3], from whose title originates the term *number theory*.

Gauss's *Disquisitiones* [4] appeared at about the same time as Legendre's book. The theoretical arithmetic of today originates from this work of Gauss. The book includes the theories of \dagger quadratic residues, \dagger quadratic forms, and **cyclotomy** (i.e., arithmetic theory of the roots of unity in the field of complex numbers), all of which appeared as well-developed theories of a remarkably high standard. The work was received with more respect than comprehension. Dirichlet made a lifelong effort to popularize the *Disquisitiones*; he also applied analytical methods to compute the \dagger class number of quadratic forms, thus giving number theory a new direction, the analytic theory of numbers. Gauss treated only binary quadratic forms; Eisenstein, Minkowski, and Siegel generalized the theory to the case of n variables. The algebraic theory of numbers has its origin in Gauss's paper on biquadratic residues. (\rightarrow 14 Algebraic Number Fields; 59 Class Field Theory; 118 Diophantine Equations; 182 Geometry of Numbers; 297 Number Theory, Elementary; 347 Quadratic Fields; 430 Transcendental Numbers.)

B. Analytic Methods

Analytic methods are sometimes used to solve arithmetic problems. For example, Legendre conjectured that any arithmetic progression of integers $a, a+d, a+2d, \dots$ contains an infinite number of primes if a and d are relatively prime. The conjecture was first proved by Dirichlet in 1837; in the proof he used an $\dagger L$ -function. Recently, proofs that do not use L -functions have been obtained, although these proofs are still not purely arithmetical. Analysis is indispensable in the formulation of some arithmetic problems. For example, let $x \in \mathbb{R}$ and $\pi(x)$ be the number of primes not exceeding x . Euclid's *Elements* already give the result $\pi(x) \rightarrow \infty$ as $x \rightarrow \infty$, but to describe the behavior of $\pi(x)$ as $x \rightarrow \infty$, notions of analysis are needed. Gauss conjectured that

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x} = 1.$$

This \dagger prime number theorem was first proved in 1896, using the results of the theory of functions of a complex variable; more elementary proofs have been obtained in recent years.

Analysis is sometimes needed to solve or simplify certain problems in number theory. The branch of mathematics treating such problems is called **analytic number theory**. The question of the extent to which analysis is really needed in dealing with such problems is itself an interesting one.

In the 20th century, analytic number theory has made rapid progress. The problem of distribution of primes has been generalized to the case of algebraic number fields. †Additive number theory dealing with †Waring's problem, †Goldbach's problem, and other problems has developed and formed a new field. We also have **geometric number theory**, which deals with †lattice-point problems. (→ 4 Additive Number Theory; 123 Distribution of Prime Numbers; 242 Lattice-Point Problems; 295 Number-Theoretic Functions; 328 Partitions of Numbers; 450 Zeta Functions.)

References

- [1] C. G. Bachet de Méziriac, Problèmes plaisants et délectables qui se font par les nombres, Lyon, 1612.
- [2] L. Euler, Vollständige Anleitung zur Algebra, Petersburg, 1770.
- [3] A. M. Legendre, Essai sur la théorie des nombres, Paris, 1798.
- [4] K. F. Gauss, Disquisitiones arithmeticae, Leipzig, 1801; English translation, Yale Univ. Press, 1966.
- [5] M. B. Cantor, Vorlesungen über Geschichte der Mathematik, Teubner, 1894–1908.
- [6] L. E. Dickson, History of the theory of numbers I, II, III, Carnegie Institution of Washington, 1919–1923 (Chelsea, 1952).
- [7] P. G. L. Dirichlet, Vorlesungen über Zahlentheorie, herausgegeben und mit Zusätzen versehen von R. Dedekind, Braunschweig, fourth edition, 1894.
- [8] J. LeVeque, Reviews of papers in number theory, vols. 1–6, Amer. Math. Soc., 1974.

297 (V.2) Number Theory, Elementary

A. The Euclidean Algorithm

We denote the set of natural numbers 1, 2, 3, ... by \mathbf{N} and the set of rational integers $0, \pm 1, \pm 2, \dots$ by \mathbf{Z} . Evidently \mathbf{Z} is an ordered †commutative ring and also an †integral domain with respect to ordinary addition and multiplication. For any $a \in \mathbf{Z}$ and $b \in \mathbf{N}$, there exists a unique pair of integers q and r such that $a = qb + r$ ($0 \leq r < b$) (**division algorithm**); q is called the **quotient** and r the **remainder** of the division of a by b . When the remainder r is zero, we say that a is a **multiple** of b , b is a **divisor** of a , and a is **divisible** by b . We denote this relation by $b|a$. After a finite number of divisions $a = q_1 b + r_1$, $b = q_2 r_1 + r_2$, $r_1 = q_3 r_2 + r_3$, ... ($b > r_1 > r_2 > \dots > 0$), we reach an equa-

tion of the form $r_k = q_{k+2} r_{k+1}$. The remainder $r = r_{k+1}$, uniquely determined in this manner by a and b , is the **greatest common divisor** (**G. C. D.**) of a and b . It can be expressed as $r = ax + by$ with integers x and y . This method of obtaining the G. C. D. is called the **Euclidean algorithm**. The greatest common divisor of a and b is denoted by (a, b) . If an integer c divides both a and b , then $c|(a, b)$. When $(a, b) = 1$, we say that a and b are **relatively prime**, and there are integral solutions x, y of the equation $ax + by = 1$. Given a pair of positive integers a and b , there exists a unique positive integer c that divides any common multiple of a and b , called the **least common multiple** (**L. C. M.**) of a and b . If d and l denote the G. C. D. and L. C. M. of a and b , respectively, we have $ab = dl$.

B. Prime Numbers

An integer p is called a **prime number** if p is larger than 1 and has no positive divisors other than 1 and itself. A positive integer is called a **composite number** if it has positive divisors other than 1 and itself. The following method of selecting the prime numbers from the sequence 2, 3, 4, 5, ..., known to the Greeks, is called **Eratosthenes' sieve**: First we discard multiples of two, thus reaching the second prime number, 3. Then we discard multiples of three and reach the next prime, 5. Continuing this process, we find the sequence of primes 2, 3, 5, 7, 11, ... by sifting out all multiples.

C. Decomposition into Primes Factors

Every integer $n > 1$ can be uniquely expressed as the product of primes; the resulting decomposition can be written as $n = p^\alpha q^\beta r^\gamma \dots$ with distinct prime factors p, q, r, \dots and corresponding exponents $\alpha, \beta, \gamma, \dots$ uniquely determined by n . This theorem is called the **fundamental theorem of elementary number theory**.

D. Perfect Numbers

We denote by $\sigma(n)$ the sum of positive divisors of n (including 1 and n itself). According as $\sigma(n)$ is greater than, equal to, or less than $2n$, we call n an **abundant number**, **perfect number**, or **deficient number**. An even number is perfect if and only if it can be represented as $n = 2^{e-1} \cdot (2^e - 1)$ with prime $2^e - 1$ (L. Euler). The existence of an odd perfect number still constitutes an open question. It is well known that, if n is odd and perfect, then n must be of the

form $n = p_0^{a_0} p_1^{a_1} \dots p_t^{a_t}$, where $p_0 \equiv a_0 \equiv 1 \pmod{4}$, and a_i is even for $i > 0$. Recently it has been proved that t must be ≥ 7 (P. H. Hagsis, Jr., *Math. Comp.*, 35 (1980)). Many necessary conditions for an odd number to be perfect have been given and repeatedly improved. It has been recently proved that there exists no odd perfect number less than 10^{50} (Hagsis, *Math. Comp.*, 27 (1973)).

Two numbers m and n are said to form a pair of **amicable numbers** if $\sigma(n) - n = m$ and $\sigma(m) - m = n$ (e.g., $m = 220$ and $n = 284$). Euler listed 61 such pairs. The following two numbers m, n , both having 152 digits, constitute the largest amicable pair currently known:

$$m = 3^4 \cdot 5 \cdot 11 \cdot 5281^{19} \cdot 29 \cdot 89 (2 \cdot 1291 \cdot 5281^{19} - 1),$$

$$n = 3^4 \cdot 5 \cdot 11 \cdot 5281^{19} (2^3 \cdot 3^3 \cdot 5^2 \cdot 1291 \cdot 5281^{19} - 1)$$

(H. J. J. te Riele, *Math. Comp.*, 28 (1974)).

E. Lionnet considered the numbers n such that the product $\prod_{d|n} d$ is equal to n^2 and called such numbers **perfect numbers of the second kind**, e.g., $n = p^3$, $n = pp'$, where p and p' are unequal prime numbers. It is also known that there are numbers n such that $\sigma(n)/n$ is an integer. For example, $2^5 \cdot 3 \cdot 7$, $2^{15} \cdot 3^5 \cdot 5^2 \cdot 7^2 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 31 \cdot 43 \cdot 257$ have this property.

E. Mersenne Numbers

A number of the form $2^e - 1$, where e is a prime, is called a **Mersenne number**. For the number $2^e - 1$ to be prime, it is necessary but not sufficient that e be prime. If a Mersenne number is a prime, it is called a **Mersenne prime**. It is not known whether there are infinitely many Mersenne primes; it has been verified, however, that for $e < 44500$, there are 27 cases of e such that $2^e - 1$ is a Mersenne prime; $e = 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 21701, 23209$, and 44497. Until the 18th century, the verifications for $e \leq 31$ were done by direct calculation. For $61 \leq e \leq 127$, the "Lucas test" was utilized to execute the computation. The remaining cases were calculated by means of electronic computers. The number $2^{44497} - 1$, which has 13395 digits, is at present the largest known prime.

F. Fermat Numbers

Numbers of the form $2^{2^v} + 1$ are called **Fermat numbers**. For a number $p = 2^e + 1$ to be a prime, it is necessary that e be a power of 2.

Fermat conjectured that numbers of the form $2^{2^v} + 1$ are all primes. In fact, for $v = 0, 1, 2, 3, 4$, the corresponding $p = 3, 5, 17, 257, 65537$ are primes; however, $2^{2^5} + 1$ is divisible by 641. It is not yet known whether there exist Fermat primes other than these five primes. For further details \rightarrow [10]. Fermat numbers are closely connected with the problem of \dagger geometric construction of regular polygons.

G. Congruence

Let m be a positive integer. Two integers a and b are said to be **congruent modulo m** if their difference is divisible by m ; we denote this relation by the **congruence** $a \equiv b \pmod{m}$ (or simply by $a \equiv b \pmod{m}$) and call m the **modulus** of this congruence. Congruence modulo m is an \dagger equivalence relation, which is compatible with the ring operations and classifies \mathbb{Z} into m classes. We thus obtain the \dagger residue class ring $\mathbb{Z}/m\mathbb{Z}$ with m elements. If p is a prime, then $\mathbb{Z}/p\mathbb{Z}$ is a field which is isomorphic to the \dagger prime field with \dagger characteristic p . A complete system of representatives of the quotient set $\mathbb{Z}/m\mathbb{Z}$ is called a **complete residue system modulo m** . On the other hand, a set of $\varphi(m)$ elements n_i (φ is \dagger Euler's function) such that $n_i \not\equiv n_j \pmod{m}$ ($i \neq j$) and $(n_i, m) = 1$ is called a **reduced residue system modulo m** . The set of all residue classes represented by a reduced residue system modulo m forms a multiplicative Abelian group of order $\varphi(m)$; we denote this group by $(\mathbb{Z}/m\mathbb{Z})^*$. If $(a, m) = 1$, then $a^{\varphi(m)} \equiv 1 \pmod{m}$. If p is a prime and $(a, p) = 1$, then $a^{p-1} \equiv 1 \pmod{p}$ (**Fermat's theorem**), since $\varphi(p) = p - 1$. When m is 2, 4, p^k or $2p^k$ ($p \neq 2$; $k = 1, 2, \dots$), $(\mathbb{Z}/m\mathbb{Z})^*$ is a \dagger cyclic group, whose \dagger generator g is called a **primitive root modulo m** (\rightarrow Appendix B, Table 1). For any a prime to m and generator g of $(\mathbb{Z}/m\mathbb{Z})^*$, there exists a unique number μ ($1 \leq \mu \leq \varphi(m)$) such that $a \equiv g^\mu \pmod{m}$. We call μ the **index** of a with respect to the basis g , and write $\mu = \text{Ind}_g a$ (\rightarrow Appendix B, Table 2). The group $(\mathbb{Z}/2^k\mathbb{Z})^*$ ($k \geq 3$) is Abelian of type $(2, 2^{k-2})$, a basis of which is formed by the residue classes represented by -1 and 5. From this follows $(p-1)! \equiv -1 \pmod{p}$ (**Wilson's theorem**). Generally, if $m = \prod_{i=1}^r m_i$, $(m_i, m_j) = 1$ for $i \neq j$, then we have $\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/m_1\mathbb{Z} + \dots + \mathbb{Z}/m_r\mathbb{Z}$ (direct sum) and $(\mathbb{Z}/m\mathbb{Z})^* \cong (\mathbb{Z}/m_1\mathbb{Z})^* \times \dots \times (\mathbb{Z}/m_r\mathbb{Z})^*$ (direct product).

The congruence $ax \equiv b \pmod{m}$ with $(a, m) = d$ is solvable if and only if $d|b$, and when it is solvable, the solution is unique modulo m/d . The simultaneous congruences $x \equiv a_i \pmod{m_i}$ ($i = 1, 2, \dots, k$) are solvable if and only if $a_i \equiv a_j \pmod{(m_i, m_j)}$ ($i, j = 1, \dots, k$), and when they are solvable, the solution is unique modulo the

greatest common multiple of m_1, \dots, m_k . In particular, when $(m_i, m_j) = 1$ ($i \neq j$), the solution is unique modulo $m_1 m_2 \dots m_k$ (**Chinese remainder theorem**). If $m = p_1^{e_1} \dots p_k^{e_k}$ is the factorization of m , solving the congruence $f(x) \equiv 0 \pmod{m}$, where $f(x)$ is a polynomial with integral coefficients, can be reduced to solving $f(x) \equiv 0 \pmod{p_i^{e_i}}$ ($i = 1, \dots, k$). Also, solving a quadratic congruence can be reduced to solving a linear congruence and a congruence of the form $x^2 \equiv a \pmod{m}$.

H. Quadratic Residues

When the congruence $x^2 \equiv a \pmod{m}$, where $(a, m) = 1$, is solvable, a is said to be a **quadratic residue** modulo m ; otherwise, a is said to be a **quadratic nonresidue** modulo m . The following two conditions are necessary and sufficient for a to be a quadratic residue modulo m : (i) a is a quadratic residue with respect to each of the prime factors p ($p \neq 2$); and (ii) $a \equiv 1 \pmod{4}$ or $a \equiv 1 \pmod{8}$ according as $4|m$ or $8|m$.

Given a prime number p and integer a prime to p , the **Legendre symbol** (a/p) is by definition 1 or -1 according as a is a quadratic residue modulo p or not. The value of this symbol is determined by $a \pmod{p}$, and we have $(ab/p) = (a/p)(b/p)$. Hence the symbol determines a \dagger character of $(\mathbf{Z}/p\mathbf{Z})^*$ of order 2. Furthermore, the congruence $(a/p) \equiv a^{(p-1)/2} \pmod{p}$ holds (**Euler's criterion**).

I. Reciprocity Law

For odd primes p and q ($p \neq q$), the formulas

$$(p/q)(q/p) = (-1)^{((p-1)/2)((q-1)/2)}, \\ (-1/p) = (-1)^{(p-1)/2}, \quad (2/p) = (-1)^{(p^2-1)/8}$$

are called the **law of quadratic reciprocity of the Legendre symbol**, the **first complementary law**, and the **second complementary law**, respectively. These laws were conjectured by Euler and first proved by Gauss, who gave seven different proofs. We now have more than fifty different proofs of these laws. P. Bachmann (*Niedere Zahlentheorie I* (1902)) lists the 47 different proofs of the laws known at the time. Gauss's first proof was elementary; his second proof used quadratic forms. The latter has been reformulated utilizing the theory of quadratic fields [4]. The fourth proof used \dagger Gaussian sums, and the sixth proof used algebraic congruences with integral coefficients. His seventh proof, contained in his posthumous works, used congruences of higher degree [4]. His fourth, sixth, and seventh proofs are related to the arithmetic of \dagger cyclo-

tomic fields [4]. The third and fifth proofs are the most elementary and simple. They are based on Gauss's lemma: Let $r_1, r_2, \dots, r_{(p-1)/2}$ be the residues of divisions of $1a, 2a, \dots, (p-1)a/2$ by an odd prime p , and let n be the number of these residues that are greater than $p/2$. Then we have $(a/p) = (-1)^n$. T. Takagi gave a simplified exposition of the third proof using geometric figures (1904). The same method was rediscovered by G. Frobenius (1914).

When m is an odd integer such that $m = \pm \prod_i p_i^{e_i}$, $(m, n) = 1$, we call $(n/m) = \prod_i (n/p_i)^{e_i}$ **Jacobi's symbol**. If m has no square factor, it is a character of $(\mathbf{Z}/m\mathbf{Z})^*$. If we put $\text{sgn } m = +1$ or -1 according as $m > 0$ or $m < 0$, we have the following **law of quadratic reciprocity of Jacobi's symbol** and its **complementary laws**:

If m and n are relatively prime odd numbers, then

$$(m/n)(n/m) \\ = (-1)^{((m-1)/2)((n-1)/2) + ((\text{sgn } m-1)/2)((\text{sgn } n-1)/2)}, \\ (-1/n) = (-1)^{(n-1)/2 + (\text{sgn } n-1)/2}, \\ (2/n) = (-1)^{(n^2-1)/4},$$

where

$$n^* = (-1)^{(n-1)/2} n.$$

Furthermore, \dagger Kronecker's symbol, another generalization of the Legendre symbol, is used in the theory of quadratic number fields (\rightarrow 347 Quadratic Fields).

References

- [1] C. F. Gauss, *Werke* 2, Königliche Gesellschaft der Wissenschaften, Göttingen, 1863.
- [2] C. F. Gauss, *Sechs Beweise des Fundamentaltheorems über quadratische Reste*, Ostwald's Klassiker Nr. 122, Wilhelm Engelmann, 1901.
- [3] L. E. Dickson, *History of the theory of numbers I*, Carnegie Institution of Washington, 1919 (Chelsea, 1952).
- [4] P. G. L. Dirichlet, *Vorlesungen über Zahlentheorie*, Herausgegeben und mit Zusätzen versehen von R. Dedekind, Vieweg, fourth edition, 1894 (Chelsea, 1969).
- [5] H. Hasse, *Vorlesungen über Zahlentheorie*, Springer, 1950.
- [6] E. G. H. Landau, *Vorlesungen über Zahlentheorie I*, Hirzel, 1927 (Chelsea, 1946).
- [7] W. J. LeVeque, *Topics in number theory*, Addison-Wesley, 1956.
- [8] H. Rademacher, *Lectures on elementary number theory*, Blaisdell, 1964.
- [9] Z. I. Borevich and I. R. Shafarevich, *Number theory*, Academic Press, 1966. (Original in Russian, 1964.)

[10] W. Sierpiński, Elementary theory of numbers, Hafner, 1964.

298 (XV.6) Numerical Computation of Eigenvalues

A. General Remarks

Numerical computation of †eigenvalues and †eigenvectors of a matrix provides a basic technique for the numerical solution of various eigenvalue problems. Roughly speaking, there are two kinds of method. In methods of the first kind, one determines the †characteristic polynomial $p(\lambda) = \det(\lambda I - A)$ (where I is the unit matrix) of A (or to give an algorithm to calculate the value of $p(\lambda)$ for an arbitrary λ), then solves the algebraic equation $p(\lambda) = 0$ numerically to obtain the eigenvalues λ_μ ($\mu = 1, 2, \dots$) (\rightarrow 301 Numerical Solution of Algebraic Equations), and finally to one determines the eigenvectors x_μ by means of the equations $(\lambda_\mu I - A)x_\mu = 0$. In methods of the second kind, one obtains eigenvalues and eigenvectors directly without resorting to the solution of an algebraic equation, relying instead upon repeated application of similarity transformations. (The power method does not fit into this classification, being a different approach altogether; \rightarrow Section C.) In particular, a real symmetric or complex Hermitian matrix is reduced approximately to a diagonal matrix. In this article we denote a given square matrix of order n by $A = (a_{ij})$ ($i, j = 1, \dots, n$).

B. The Jacobi Method

The Jacobi method is an iterative technique for determining all the eigenvalues and eigenvectors of a real symmetric matrix [7]. Before the advent of high-speed computers it was not considered practical, but at present it is one of the most compact and elegant methods. The following algorithm can be extended to Hermitian matrices by replacing †orthogonal transformations by suitable †unitary transformations.

Roughly, the method transforms a given matrix $A = (a_{ij})$ ($a_{ji} = a_{ij}$; $i, j = 1, \dots, n$) into a diagonal one by repeated application of 2-dimensional rotations of the reference axes. We first put $A^{(0)} = A$, $U^{(0)} = I$, and compute $A^{(1)}$, $A^{(2)}$, ..., $U^{(1)}$, $U^{(2)}$, ... successively as follows:

(1) Select an off-diagonal element of $A^{(0)} =$

($a_{ij}^{(0)}$) with the maximum absolute value and denote it by $a_{pq}^{(0)}$.

(2) Compute

$$\tan \theta = \frac{2a_{pq}^{(0)} \operatorname{sgn}(a_{pp}^{(0)} - a_{qq}^{(0)})}{|a_{pp}^{(0)} - a_{qq}^{(0)}| + \sqrt{(a_{pp}^{(0)} - a_{qq}^{(0)})^2 + 4(a_{pq}^{(0)})^2}}$$

(where $\operatorname{sgn} x = 1, 0$, or -1 according as $x > 0$,
 $= 0$, or < 0),

$$\cos \theta = (1 + \tan^2 \theta)^{-1/2}, \quad \text{and}$$

$$\sin \theta = \tan \theta \cdot \cos \theta;$$

and form $T^{(0)} = (t_{ij}^{(0)})$, where $t_{pp}^{(0)} = t_{qq}^{(0)} = \cos \theta$, $t_{ii}^{(0)} = 1$ for $i \neq p, q$, $-t_{pq}^{(0)} = t_{qp}^{(0)} = \sin \theta$, and $t_{ij}^{(0)} = 0$ for all other (i, j) .

(3) Determine $A^{(u+1)}$ and $U^{(u+1)}$ by $A^{(u+1)} = T^{(0)'} A^{(0)} T^{(0)}$ ($'$ denotes the transpose) and $U^{(u+1)} = U^{(0)} T^{(0)}$. In this process, $T^{(0)}$ represents an orthogonal transformation (rotation) in the plane spanned by the p th and q th coordinate axes such that $a_{pq}^{(u+1)} = a_{pq}^{(u+1)}$ is nullified. If we put $N(B) = \sum_{i,j} b_{ij}^2$ and $M(B) = \sum_{i \neq j} b_{ij}^2$, then $N(B)$ is invariant under an orthogonal transformation, so that $N(A^{(0)}) = N(A)$. Furthermore, since $a_{ii}^{(u+1)} = a_{ii}^{(0)}$ ($i \neq p, q$) and $(a_{pp}^{(u+1)})^2 + (a_{qq}^{(u+1)})^2 = (a_{pp}^{(0)})^2 + (a_{qq}^{(0)})^2 + 2(a_{pq}^{(0)})^2$, we have $M(A^{(u+1)}) = M(A^{(0)}) - 2(a_{pq}^{(0)})^2$. Since $a_{pq}^{(0)}$ has the maximum absolute value among all the off-diagonal elements, we have $(a_{pq}^{(0)})^2 \geq M(A^{(0)})/(n^2 - n)$. Therefore

$$\begin{aligned} M(A^{(u+1)}) &\leq (1 - 2/(n^2 - n)) M(A^{(0)}) \\ &\leq (1 - 2/(n^2 - n))^{u+1} M(A) \\ &< M(A) \exp(-2(u+1)/(n^2 - n)). \end{aligned}$$

It has been proved a fortiori [13] that, after $M(A^{(0)})$ comes down to below a certain threshold value, the convergence of the iteration process becomes quadratic, i.e., there is a number c determined by the order n of A and the arrangement of the eigenvalues of A such that $M(A^{(u+n(n-1)/2)}) < c(M(A^{(0)}))^2$. Since the set of eigenvalues of $A^{(0)}$ coincides with that of A and the eigenvalues of an arbitrary symmetric matrix B can be made to correspond in a one-to-one way to its diagonal elements in such a way that the difference between an eigenvalue and the corresponding diagonal element is not greater than $M(B)^{1/2}$, $a_{ii}^{(0)}$ tends to λ_i ($i = 1, \dots, n$) as l tends to infinity. Moreover, as l tends to infinity, each column vector of $U^{(0)} = (u_{ij}^{(0)})$ tends to the corresponding eigenvector, in the sense that $\sum_{k=1}^n a_{ik} u_{kj}^{(0)} - \lambda_j u_j^{(0)} \rightarrow 0$.

The number of arithmetic computations required to obtain $A^{(u+1)}$ and $U^{(u+1)}$ from $A^{(0)}$ and $U^{(0)}$ is at most proportional to n , so that for a given $\varepsilon > 0$, the arithmetic required to reduce $\max_i |a_{ii}^{(0)} - \lambda_i|$ below $\varepsilon(M(A))^{1/2}$ is at most proportional to n^3 (because l is at most

proportional to n^2). On the other hand, the search for an off-diagonal element of $A^{(i)}$ with the maximum absolute value, if it is done by simply comparing all the elements, will require effort proportional to n^2 , so that the amount of work required by the searching process is proportional to n^4 . To bypass this searching process, the **cyclic Jacobi method** and the **threshold Jacobi method** are often used. The former method adopts as $a_{pq}^{(i)}$ the off-diagonal element for which $q > p$ and $i = (p-1)(n-p/2) + (q-p)$, that is, $a_{12}^{(1)}, a_{13}^{(2)}, \dots, a_{1n}^{(n-1)}, a_{23}^{(n)}, a_{24}^{(n+1)}, \dots$ are adopted in this sequence. The latter method adopts as $a_{pq}^{(i)}$ off-diagonal elements in a sequence similar to the one above as long as they exceed a given threshold value; but if an element is less than that threshold value, then the next element in the sequence is a candidate for adoption as $a_{pq}^{(i)}$, where the threshold value is made to decrease gradually as the iteration process proceeds. However, the search for an element with the maximum absolute value can be done more effectively by taking account of the fact that only the matrix elements lying in rows p and q and in columns p and q change their values when we transform $A^{(i)}$ into $A^{(i+1)}$. In fact, we can record for each row the value as well as the position of the (off-diagonal) element with the maximum absolute value in that row. By so doing, the effort of searching for an off-diagonal element with the maximum absolute value can be reduced to something proportional to n on the average.

C. The Power Method

The **power method** is suitable for obtaining only the eigenvalue of maximum absolute value [6]. Let us assume that the eigenvalues $\lambda_1, \dots, \lambda_n$ of A are arranged so that $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$, with λ_1 real, and denote by y_1 the left eigenvector corresponding to λ_1 (which means $y_1(\lambda_1 I - A) = 0$). Starting from an arbitrary (real) vector $x^{(0)}$ such that $(y_1, x^{(0)}) \neq 0$ and $x_{i_0}^{(0)} = 1$ for a prescribed i_0 , we compute $\theta^{(0)}, \theta^{(1)}, \dots$ and $x^{(1)}, x^{(2)}, \dots$ by $Ax^{(i)} = \theta^{(i)} x^{(i+1)}$ ($i = 0, 1, 2, \dots; x_{i_0}^{(i+1)} = 1$). Then we have $\lim_{i \rightarrow \infty} \theta^{(i)} = \lambda_1$ and $\lim_{i \rightarrow \infty} x^{(i)} = x_1$ (the eigenvector corresponding to the eigenvalue λ_1). The rate of convergence depends on $|\lambda_1/\lambda_2|$ if the elementary divisor of A corresponding to λ_1 is linear, but in the case of a nonlinear elementary divisor, the convergence is too slow for practical purposes. If $\lambda_1 \neq \lambda_2$ and $|\lambda_1| = |\lambda_2| > |\lambda_3| \geq \dots \geq |\lambda_n|$, then the sequences of $\theta^{(i)}$ and $x^{(i)}$ computed by the formulas above do not converge but in general oscillate. However, from $\theta^{(i)}, \theta^{(i+1)}, x^{(i)}, x^{(i+1)}, x^{(i+2)}$ for a sufficiently large i , we can obtain

approximate eigenvalues λ_1 and λ_2 as the two roots of the quadratic equation in λ :

$$\begin{vmatrix} \chi_i^{(i)} & \chi_j^{(i)} & \theta^{(i)} \theta^{(i+1)} \\ \chi_i^{(i+1)} & \chi_j^{(i+1)} & \theta^{(i)} \lambda \\ \chi_i^{(i+2)} & \chi_j^{(i+2)} & \lambda^2 \end{vmatrix} = 0$$

(i and j arbitrary, $i \neq j$).

The corresponding eigenvectors are given by $x_1 = \lambda_2 x^{(i+1)} - x^{(i+2)}$ and $x_2 = \lambda_1 x^{(i+1)} - x^{(i+2)}$. Complex conjugate pairs of eigenvalues can be dealt with in this manner. This is useful also when $|\lambda_1| = |\lambda_2|$. The extension to the case of more than two eigenvalues with the same maximum absolute value is obvious.

In order to determine the remaining eigenvalues by the power method we have to combine it with the deflation or transformation of matrices, as discussed later in this section and in the following section. The amount of computation depends on the arrangement of the eigenvalues of A and the required accuracy. We note that the multiplication of a matrix by a vector requires an amount of computation proportional to n^2 .

Improvement in the approximations can be incorporated into the power method. If $v = x_\mu + O(\epsilon)$ is an approximation to the eigenvector x_μ corresponding to an eigenvalue λ_μ of a real symmetric matrix A , then the Rayleigh quotient $\lambda_R = (v, Av)/(v, v)$ affords a good approximation to λ_μ . In fact, we have $|\lambda_\mu - \lambda_R| = O(\epsilon^2)$.

If λ_i ($i = 1, \dots, n$) is an eigenvalue of A and x_i the corresponding eigenvector, then $P(\lambda_i)$ ($i = 1, \dots, n$) is an eigenvalue of $P(A)$ and x_i the corresponding eigenvector, where $P(\xi)$ is a polynomial in ξ and ξ^{-1} . This fact can be utilized to transform the magnitudes of eigenvalues to accelerate the convergence of the power method, to separate eigenvalues with the same absolute value, or to obtain intermediate eigenvalues. The choice $P(\lambda) = \lambda - c$ or $(\lambda - c)^{-1}$ is particularly useful, where c is an appropriate constant. In fact, an efficient algorithm known as **inverse iteration** exists for computing an approximate eigenvector when a good approximation to an eigenvalue is known. Given a trial eigenvector x corresponding to a computed eigenvalue c , one computes an improved approximate eigenvector y by solving $(A - cI)y = x$ or $y = (A - cI)^{-1}x$.

Aitken's δ^2 -method is also efficient in accelerating the convergence of the power method.

When an eigenvalue-eigenvector pair is known, another eigenvalue-eigenvector pair can be computed using the process known as **deflation**. If an eigenvalue λ_μ and the corresponding eigenvector x_μ (and also the corresponding left eigenvector y_μ if necessary) of A are known, it is possible to "subtract" them

from A to get a problem containing only the remaining eigenvalues. Such deflations are often used in combination with the power method. The following are two examples of deflation methods.

(1) Assuming that x_μ and y_μ are normalized in such a way that $(y_\mu, x_\mu) = 1$, form $B = A - \lambda_\mu x_\mu y_\mu'$. Then B has the same set of eigenvalues and eigenvectors as A except for λ_μ . The eigenvalue and the eigenvector of B corresponding to λ_μ of A are 0 and x_μ , respectively. This kind of deflation process can be generalized to the case of nonlinear elementary divisors, but that becomes somewhat more complicated.

(2) After normalizing x_μ so that its n th component $x_{\mu n}$ is equal to 1, form $B = (b_{ij})$; $b_{ij} = a_{ij} - x_{\mu i} a_{nj}$ ($i, j = 1, \dots, n-1$). Then B has the same set of eigenvalues as A except for λ_μ . If w_k is the eigenvector corresponding to the eigenvalue λ_k of B , then the corresponding eigenvector x_k of A is given by $x_{ki} = w_{ki} + d_k x_{\mu i}$ ($i = 1, \dots, n-1$), $x_{kn} = d_k$, where d_k is determined from $\sum_{i=1}^{n-1} n_i w_{ki} = (\lambda_k - \lambda_\mu) d_k$ if $\lambda_k \neq \lambda_\mu$. If $\lambda_k = \lambda_\mu$ and $r \equiv \sum_{i=1}^{n-1} a_{ni} w_{ki} = 0$, then we can put $d_k = 0$. If $\lambda_k = \lambda_\mu$ and $r \neq 0$, A has a nonlinear elementary divisor for $\lambda_k = \lambda_\mu$, and the x_k defined by $x_{ki} = w_{ki}/r$ and $x_{kn} = 0$ is a generalized eigenvector of A in the sense that $Ax_k = \lambda_\mu x_k + x_\mu$.

D. Transformation of Matrices

There are a number of methods of transforming a given matrix A by means of a suitable similarity transformation $A \rightarrow B = S^{-1}AS$ into another matrix B for which it is easier to solve the eigenvalue problem. The **Givens method** [9] transforms a symmetric matrix A into a **tridiagonal matrix** B (i.e., a matrix such that $b_{ij} = 0$ for $|i-j| \geq 2$) by means of a matrix S which is the product of 2-dimensional rotation matrices. The **Householder method** [10] also transforms a symmetric A into a tridiagonal B by means of an orthogonal matrix S of special type, i.e., a reflection matrix $I - 2uu^*$ ($u^*u = 1$, $u^* =$ conjugate transpose of u), and the **Lanczos method** [8] transforms a general A into a tridiagonal B . To general matrices the following methods are also applicable: (1) The **Danilevskii method** [1] transforms A into its companion matrix B by repeated application of elimination operations. Here, row interchanges can be combined to increase numerical stability [14]. (2) The **Hessenberg method** [3] transforms A into a B such that $b_{ij} = 0$ for $i-j \geq 2$ with a triangular S . (3) The Givens method [9] transforms A into a B of the same form as in (2) by repeated application of 2-dimensional rotations. This method now tends to give way to the Householder method or to

the elimination method with row interchange. All these methods require an amount of computation proportional to n^3 . In general, special treatment is necessary for the case of multiple eigenvalues. As an example, we explain the Givens method for general matrices. Let $N = (n-1)(n-2)/2$. For $i = 0, 1, \dots, N-1$, choose $(p, q) = (3, 2), (4, 2), \dots, (n, 2); (4, 3), (5, 3), \dots, (n, 3); \dots; (n-1, n-2), (n, n-2); (n, n-1)$ in this order. Using $T^{(i)}$ of the same form as in the Jacobi method, and setting $\tan \theta = a_{p,q-1}^{(i)}/a_{q,q-1}^{(i)}$ in this case, calculate $A^{(0)} = A$, $U^{(0)} = I$, $A^{(i+1)} = T^{(i)'} A^{(i)} T^{(i)}$, $U^{(i+1)} = U^{(i)} T^{(i)}$ ($i = 0, 1, \dots, N-1$), and put $B = A^{(N)}$. Then we have $b_{ij} = 0$ for $i-j \geq 2$. We can solve the eigenvalue problem for this simplified B and then retransform the eigenvectors of B thus obtained into those of A by means of $U^{(N)}$. It should be noted that the method of bisection based on Sturm's theorem is effectively used to solve the characteristic equation of a real tridiagonal matrix [14]. This method is remarkably stable numerically. It is used when the number of eigenvalues to be computed is small relative to the order of the given matrix. If all eigenvalues are desired, an alternative method such as the QR method (\rightarrow Section F) is recommended.

E. The Lanczos Method

A more detailed exposition of the Lanczos method is now given. Let A be a given real matrix of order n . Pick two vectors c_1 and \tilde{c}_1 . Determine recursively the vectors c_{i+1} and \tilde{c}_{i+1} , $i = 1, \dots, n$, that satisfy the following conditions: (i) $\gamma_{i+1} c_{i+1} = A c_i - \alpha_i c_i - \beta_i c_{i-1} \equiv b_{i+1}$, $\tilde{\gamma}_{i+1} \tilde{c}_{i+1} = A' \tilde{c}_i - \tilde{\alpha}_i \tilde{c}_i - \tilde{\beta}_i \tilde{c}_{i-1} \equiv \tilde{b}_{i+1}$, $i = 2, \dots, n$; (ii) c_{i+1} is orthogonal to \tilde{c}_{i-1} and \tilde{c}_i ; (iii) \tilde{c}_{i+1} is orthogonal to c_{i-1} and c_i , where $\alpha_i, \beta_i, \tilde{\alpha}_i$, and $\tilde{\beta}_i$ are scalars, and where γ_i and $\tilde{\gamma}_i$ are normalizing scalars. Actually, $\alpha_i = \tilde{c}_i' A c_i / \tilde{c}_i' c_i = \tilde{\alpha}_i$, $\beta_i = \gamma_i \tilde{c}_i' \tilde{c}_i / \tilde{c}_{i-1}' c_{i-1}$, $\tilde{\beta}_i = \gamma_i \tilde{c}_i' c_i / c_{i-1}' \tilde{c}_{i-1}$, where $c_i' \tilde{c}_i$, $i = 1, \dots, n$, are assumed nonzero. It can be shown that c_{i+1} is orthogonal to every \tilde{c}_j , $1 \leq j \leq i$, and that \tilde{c}_{i+1} is orthogonal to every c_j , $1 \leq j \leq i$ ($c_{n+1} = \tilde{c}_{n+1} = 0$). The conclusion is that the given matrix A is similar to the tridiagonal matrix $H = (h_{ij})$, where $h_{ii} = \alpha_i$ ($i = 1, \dots, n$), $h_{i,i+1} = \beta_{i+1}$, $h_{i+1,i} = \tilde{\gamma}_{i+1}$ ($i = 2, \dots, n$). In fact, if C denotes the matrix whose j th column equals c_j ($j = 1, \dots, n$), one obtains $C^{-1}AC = H$. Thus the eigenvalues of A are identical to those of H .

In principle, the Lanczos method transforms a given matrix to a tridiagonal matrix similar to it if every $c_i' \tilde{c}_i$ is nonzero ($i = 1, \dots, n$). If $c_i' \tilde{c}_i$ does vanish for a certain i , one selects c_1 and \tilde{c}_1 again and restarts. On the other hand, if $b_{i+1} = 0$ one chooses an arbitrary vector c_{i+1} orthogonal to $\tilde{c}_1, \dots, \tilde{c}_i$ and if $\tilde{b}_{i+1} = 0$ one selects an arbitrary vector \tilde{c}_{i+1} orthogonal to

c_1, \dots, c_i . In the actual numerical computation, the distinction between zero vectors and non-zero vectors is usually blurred by rounding errors.

In the application of the Lanczos method one often observes the loss of orthogonality $c_i^T \tilde{c}_j = 0$ ($i \neq j$). This usually results from cancellation errors in the computation of b_{i+1} and \tilde{b}_{i+1} . If the orthogonality is lost, $C^{-1}AC$ may significantly deviate from a tridiagonal matrix. As a practical means for preserving the indicated orthogonality one can reorthogonalize the vectors c_i and \tilde{c}_i . Indeed, one can add to the computed c_{i+1} a linear combination of c_1, \dots, c_i so that the sum is orthogonal to $\tilde{c}_1, \dots, \tilde{c}_i$, then take the sum as c_{i+1} after properly normalizing. A similar process applies to \tilde{c}_{i+1} . The Lanczos method is often applied in double precision. It has been suggested that one first reduces the given matrix to an upper Hessenberg form by using the Hessenberg method with row interchange before applying the Lanczos method [15].

A further remark on the Lanczos method is in order. Recall that γ_i is determined from $Ac_i, c_{i-1}, \alpha_{i-1}$, and β_{i-1} ; β_i from $c_i, \tilde{c}_i, c_{i-1}, \tilde{c}_{i-1}$, and $\tilde{\gamma}_i$; and α_i from Ac_i, c_i , and \tilde{c}_i . This shows that the Lanczos method applied to a sparse matrix (a matrix whose elements are almost all zero) requires only a memory proportional to n . The Householder method, on the other hand, requires memory proportional to n^2 even for a sparse matrix.

For a real symmetric matrix one can modify the Lanczos method for a general matrix so that C is orthogonal and H is real, symmetric, and tridiagonal (the details are omitted).

F. The QR Method

The QR method was discovered independently by J. G. F. Francis and by V. N. Kublanovskaya in 1961 [15]. The method has been improved and extended considerably since then. In the usual application to matrix eigenvalue problems, the QR method provides the most efficient iterative process for finding all eigenvalues of a given matrix. The matrix is reduced by means of a similarity transformation to Hessenberg form or to tridiagonal form before application of the QR method; the reduction process may be effected by the Householder method or by the elimination method with row interchange. (If the given matrix is a complex matrix, the latter is preferable.) The reason is that one step of the QR process applied to a full matrix is prohibitively expensive, requiring a number of operations propor-

tional to n^3 , while one step of the QR process applied to a Hessenberg matrix requires a number of operations proportional to n^2 .

We now describe one of the most useful versions of the QR method. Let $A = A_0$ be a given matrix, and define a sequence of matrices A_1, A_2, \dots obtained from A_0 as follows. At the i th ($i = 0, 1, \dots$) step, choose an appropriate constant s_i , called an origin shift, according to the process described below, and decompose $A_i - s_i I$ into the product of a unitary matrix Q_i and an upper Hessenberg matrix R_i : $A_i - s_i I = Q_i R_i$. The matrix A_{i+1} is then defined by $A_{i+1} = R_i Q_i + s_i I = Q_i^* A_i Q_i$. Hence A_{i+1} is similar to A_i . The QR method is closely related to the power method and to the inverse power method. We describe this relationship in order to obtain an insight into the nature of the QR method. To this end, we first state a well-known theorem: Let A be diagonalizable, and let its eigenvalues λ_i ($i = 1, \dots, n$) have distinct moduli, say $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$. Let $X^{-1}AX = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, and let X have an LU-decomposition $X = LU$, where L and U are, respectively, lower triangular matrix and an upper triangular matrix. Then the QR algorithm without an origin shift ($s_i = 0, i = 0, 1, 2, \dots$) behaves as follows: $a_{pq}^{(i)} \rightarrow 0$ ($p > q$); $a_{pp}^{(i)} \rightarrow \lambda_p$; $a_{pq}^{(i)}$ oscillates ($p < q$), $p, q = 1, 2, \dots, n, i \rightarrow \infty$, where $a_{pq}^{(i)}$ denotes the (p, q) th element of A_i . In other words, A_i approaches an upper triangular matrix as $i \rightarrow \infty$ so that the diagonal elements of A_i converge to the eigenvalues of A .

The relationship between the QR algorithm and the power method is given by $(Q_0 Q_1 \dots Q_i)(R_i R_{i-1} \dots R_0) = (A - s_0 I)(A - s_1 I) \dots (A - s_i I)$. If $s_0 = s_1 = \dots = s_i = 0$, then the right-hand side reduces to A^{i+1} , and since the product $R_i \dots R_0$ is upper triangular, the first column of $\tilde{Q}_i \equiv Q_0 Q_1 \dots Q_i$ equals a scalar multiple of $A^{i+1} e_1$, where $e_1 = (1, 0, 0, \dots, 0)^T$. Therefore, by the power method, the first column of \tilde{Q}_i converges to an eigenvector corresponding to the eigenvalue λ_1 of A that has the largest modulus, under a certain fairly mild condition. Since $A_{i+1} = \tilde{Q}_i^* A \tilde{Q}_i$, $A_{i+1} e_1 = \tilde{Q}_i^* A \tilde{Q}_i e_1 \approx \lambda_1 e_1$ for large i . In other words, the first column of A_i converges to the vector $(\lambda_1, 0, 0, \dots, 0)^T$ as $i \rightarrow \infty$.

The relationship of the QR method to the inverse power method and to the Rayleigh quotient is now explained. From the definition of the QR algorithm with an origin shift, we have $Q_i = [(A - s_i I)^{-1}]^* R_i^*$. Since R_i^* is a lower triangular matrix, the n th column of Q_i equals $Q_i e_n = [(A_i - s_i I)^{-1}]^* R_i^* e_n = (A - s_i I)^{-1} \cdot$ (a scalar multiple of e_n), where $e_n = (0, 0, \dots, 0, 1)^T$. This last process of obtaining the last column of Q_i represents a process known as the inverse power method. If s_i is close to an eigenvalue of

A_i (and hence to an eigenvalue of A), the last column of A_i gives a good approximate eigenvector of A corresponding to the eigenvalue under consideration. Now, if x is a given column vector such that $x^*x = 1$, the value of λ which minimizes $(Ax - \lambda x)^*(Ax - \lambda x)$ is given by the Rayleigh quotient x^*Ax . If $x(x^*x = 1)$ happens to be an eigenvector of A , the Rayleigh quotient equals an eigenvalue of A . If we take $x = e_n$, the corresponding Rayleigh quotient equals a_{nn} . Therefore if we take the (n, n) th element of A_i as the origin shift s_i , then s_i can be regarded as the best approximation to an eigenvalue of A_i (hence of A) when e_n is taken as an approximate eigenvector of A_i , in the sense that $\lambda = s_i$ minimizes the functional $(A_i e_n - \lambda e_n)^*(A_i e_n - \lambda e_n)$.

Under the same condition as in the preceding theorem, the rate of convergence of the QR method with an origin shift is given as follows: $a_{pq}^{(i)} (n \geq p \geq q \geq 1)$ is asymptotically proportional to $(\lambda_p/\lambda_q)^i$ for $s_i = 0$ ($i = 0, 1, \dots$) (no origin shift); and with the origin shift s_i , the behavior of $a_{pq}^{(i)}$ at the i th step is determined by $(\lambda_p - s_i)/(\lambda_q - s_i)$. If $s_i = \lambda_n$ (= the eigenvalue of A with the least modulus), each element in the n th row of A_{i+1} except $a_{nn}^{(i+1)}$ exhibits rapid decrease in modulus. A natural and practical choice of the origin shift s_i is given as follows: (i) for a real tridiagonal matrix, s_i is taken to be that eigenvalue of the 2×2 matrix situated at the lower right corner of A_i that is closer to $a_{nn}^{(i)}$ [14]; (ii) for a real upper Hessenberg matrix, the two eigenvalues of the 2×2 matrix situated at the lower right corner of A_i as s_i and s_{i+1} [14]; (iii) for a complex upper Hessenberg matrix, s_i is chosen as in (i) [17, COMQR].

The QR method with origin shift would eventually make each element in the n th row, except $a_{nn}^{(i)}$, smaller in modulus than a prescribed positive number. At this stage of iteration, $a_{nn}^{(i)}$ is taken as an approximate eigenvalue of A . The QR method is then applied anew to the $(n-1) \times (n-1)$ matrix obtained from A_i by deleting the n th row and the n th column of A_i , and another approximate eigenvalue of A is obtained. The method proceeds similarly until all the eigenvalues of A are computed. For maximum accuracy the eigenvalues of the given matrix should be computed in the order of increasing modulus. A word of caution is in order. When the given matrix A has elements of greatly varying modulus, rearrangement of elements of A may be necessary before applying the QR method with an explicit origin shift, where $A_i - s_i I$ ($i = 0, 1, \dots$) is explicitly computed.

The reader is referred to [14–16] for details of the QR method.

G. Generalized Eigenvalue Problem $Ax = \lambda Bx$

An eigenvalue problem of the type $Ax = \lambda Bx$ is called a **generalized eigenvalue problem** and is often encountered in applied mathematics. A necessary and sufficient condition for λ to be an eigenvalue is $\det(A - \lambda B) = 0$. If B^{-1} exists, then $Ax = \lambda Bx$ is equivalent to $B^{-1}Ax = \lambda x$ and hence has n eigenvalues, where n is the order of A . If B^{-1} does not exist, the eigenvalue problem $Ax = \lambda Bx$ has at most $n-1$ eigenvalues.

We restrict ourselves to one of the most important cases: that where A is real and symmetric and B is real, symmetric, and positive definite. In this case one could solve the problem by reformulating it as $B^{-1}Ax = \lambda x$, where $B^{-1}A$ is explicitly computed. However, $B^{-1}A$ is not in general symmetric. Moreover, when B has eigenvalues of widely different moduli, the elements of B^{-1} may have widely different moduli as well, which would in turn make the computation of those eigenvalues of $B^{-1}A$ of smaller moduli difficult. An efficient and numerically stable method is known which obviates the aforementioned difficulty by exploiting the symmetry of A and B . An outline of this procedure is now given. Since B is positive definite, a lower triangular matrix L exists such that $LL' = B$. This is called the Cholesky decomposition of B . The elements of L can be computed by equating the corresponding elements in $LL' = B$. The eigenvalue problem $Ax = \lambda Bx$ is now equivalent to $L^{-1}A(L')^{-1}y = \lambda y$ with $y = L'x$, where the matrix $L^{-1}A(L')^{-1} = P$ is computed in two stages by solving $LX = A$ and $PL' = X$. In this last equation it is enough to compute the upper right half of P , because P is symmetric while X is not generally symmetric. The upper right half of P can be computed by equating the corresponding elements in $PL' = X$. The eigenvalues λ of $Ax = \lambda Bx$ are given by the eigenvalues of the matrix P . Other types of eigenvalue problems, such as $y'AB = \lambda y'$, $BAy = \lambda y$, and $x'BA = \lambda x'$, often appear in practice, where A is real and symmetric and B is real, symmetric, and positive definite. By using the Cholesky decomposition of B , one can reduce any one of these eigenvalue problems to the ordinary eigenvalue problem for a real symmetric matrix [4].

If A and B are general matrices in $Ax = \lambda Bx$, the following method is known to be effective [18]. First, reduce B to an upper triangular matrix by applying $n-1$ Householder transformations from the left. This reduces the eigenvalue problem $Ax = \lambda Bx$ to the case where B is upper triangular. Next, apply to A a sequence of plane rotations of Givens type from the left, thereby reducing the eigenvalue problem $Ax = \lambda Bx$ to the case where A is an

Numerical Computation of Eigenvalues

upper Hessenberg matrix and B is an upper triangular matrix. Then apply the QR method to $B^{-1}Ax = \lambda x$ without explicitly computing $B^{-1}A$ to reduce the eigenvalue problem $Ax = \lambda Bx$ to the case where A and B are both approximately upper triangular. The eigenvalues of $Ax = \lambda Bx$ are then easily computed as ratios of corresponding diagonal elements. This method is called the **QZ method** [18].

A collection of about 50 excellent FORTRAN subroutines for various types of matrix eigenvalue problems is contained in [18]. These subroutines are in most part translations from ALGOL procedures given in [14].

References

- [1] H. Wayland, Expansion of determinantal equations into polynomial form, *Quart. Appl. Math.*, 2 (1945), 277–306.
- [2] P. S. Dwyer, *Linear computation*, Wiley, 1951.
- [3] R. Zurmühl, *Matrizen*, Springer, 1950.
- [4] L. Collatz, *Eigenwertaufgaben mit technischen Anwendungen*, Akademische Verlag, 1949.
- [5] R. V. Mises and H. Geiringer, *Praktische Verfahren der Gleichungsauflösung*, *Z. Angew. Math. Mech.*, 9 (1929), 58–77, 152–164.
- [6] E. Bodewig, *Matrix calculus*, Interscience, 1956.
- [7] C. G. J. Jacobi, Über ein leichtes Verfahren, die in der Theorie der Säkularstörungen vorkommenden Gleichungen numerisch aufzulösen, *J. Reine Angew. Math.*, 30 (1846), 51–95.
- [8] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards*, 45 (1950), 255–282.
- [9] W. Givens, Computation of plane unitary rotations transforming a general matrix to triangular form, *SIAM J.*, 6 (1958), 26–50.
- [10] J. H. Wilkinson, Householder's method for symmetric matrices, *Numer. Math.*, 4 (1963), 354–361.
- [11] P. A. White, The computation of eigenvalues and eigenvectors of a matrix, *SIAM J.*, 6 (1958), 393–437.
- [12] S. H. Crandall, *Engineering analysis*, McGraw-Hill, 1956.
- [13] A. Schönage, Zur quadratischen Konvergenz des Jacobi-Verfahrens, *Numer. Math.*, 6 (1964), 410–412.
- [14] J. H. Wilkinson and C. Reinsch, *Handbook for automatic computation II, Linear algebra*, Springer, 1971.
- [15] J. H. Wilkinson, *The algebraic eigenvalue problem*, Clarendon Press, 1965.
- [16] G. W. Stewart, *Introduction to matrix computations*, Academic Press, 1973.
- [17] B. T. Smith, J. M. Boyle, J. J. Dongarra, B. S. Garbow, Y. Ikebe, V. C. Klema, and C. B. Moler, *Matrix eigensystem routines—EISPACK guide*, Springer, second edition, 1976.
- [18] C. B. Moler and G. W. Stewart, An algorithm for generalized matrix eigenvalue problems, *SIAM J. Numer. Anal.*, 10 (1973), 241–256.

299 (XV.7) Numerical Integration

A. Interpolatory Integration Formulas

Numerical integration is a method of finding an approximate numerical value of a definite integral of a given function $f(x)$. Usually the integral $\int_a^b f(x)dx$ or $\int_a^b f(x)w(x)dx$ ($w(x)$ is the **weight function**) is approximated by a linear combination $\sum_{i=1}^n W_i f(x_i)$ of the values of the integrand at the points x_1, x_2, \dots, x_n . Integration formulas are divided into two groups, the interpolatory formulas and the formulas based on variable transformation.

In order to obtain an **interpolatory formula**, we interpolate over the integrand $f(x)$ at n points x_1, x_2, \dots, x_n by means of the † Lagrange interpolation polynomial of degree $n-1$, and then integrate the polynomial over $[a, b]$. Depending on the selection of the points x_i and the weights W_i , we have several kinds of formulas.

(1) Newton-Cotes Formulas. We assume $w(x) = \text{constant}$ and $x_i = x_0 + ih$ ($i = 0, 1, \dots, n$). The weights W_i are so determined that the value of the integral can be calculated accurately if the integrand $f(x)$ is a polynomial whose degree does not exceed n . There formulas are called the **Newton-Cotes formulas**: $\int_{x_0}^{x_1} f(x)dx = (f_0 + f_1)h/2$ for $n = 1$ (**trapezoidal rule**), $\int_{x_0}^{x_2} f(x)dx = (f_0 + 4f_1 + f_2)h/3$ for $n = 2$ (**Simpson's 1/3 rule**), and $\int_{x_0}^{x_3} f(x)dx = (f_0 + 3f_1 + 3f_2 + f_3)3h/8$ for $n = 3$ (**Simpson's 3/8 rule**). The truncation errors of these three formulas are given by $h^3 f^{(2)}(\xi)/12$, $h^5 f^{(4)}(\xi)/90$, $3h^5 f^{(4)}(\xi)/80$, respectively, where ξ is a number in the interval of integration and $f^{(i)}$ denotes the i th derivative of f (differentiability is assumed). For an even n , the polynomial of degree $n+1$ can also be integrated accurately by these formulas.

In Table 1 the coefficients A and B_i of the Newton-Cotes formulas $hA \sum_{i=0}^n B_i f(x_i)$, $h = (x_n - x_0)/n$, and the coefficients C of the error

Table 1

<i>n</i>	<i>A</i>	<i>B</i> ₀	<i>B</i> ₁	<i>B</i> ₂	<i>B</i> ₃	<i>B</i> ₄	<i>B</i> ₅	<i>B</i> ₆	<i>B</i> ₇	<i>B</i> ₈	<i>C</i>
1	1/2	1	1								−1/12
2	1/3	1	4	1							−1/90
3	3/8	1	3	3	1						−3/80
4	4/45	7	32	12	32	7					−8/945
5	5/288	19	75	50	50	75	19				−275/12096
6	1/140	41	216	27	272	27	216	41			−9/1400
7	7/17280	751	3577	1323	2989	2989	1323	3577	751		−8183/518400
8	4/14175	989	5888	−928	10496	−4540	10496	−928	5888	989	−2368/467775

term $Ch^{p+1}f^{(p)}(\xi)$, where $p = n + 2$ if n is even and $p = n + 1$ if n is odd, are given.

When the interval $[a, b]$ of integration is large, we usually divide it into small subintervals and apply formulas for small n for each part rather than formulas for large n for the whole interval. For example, if the interval is divided into m equal subintervals, we get the following formula by applying the trapezoidal rule for each subinterval:

$$\int_a^b f(x)dx = h((f_0 + f_m)/2 + (f_1 + f_2 + \dots + f_{m-1})),$$

where $x_0 = a$, $x_m = b$, $h = (b - a)/m$, with truncation error $(b - a)^3 f^{(2)}/(12m^2)$. (Here, as in the rest of the article, $f^{(i)}$ stands for $f^{(i)}(\xi)$, with differentiability assumed as before.) By applying Simpson's 1/3 rule we obtain the formula

$$\begin{aligned} \int_a^b f(x)dx \\ = \frac{2}{3}h((f_0 + f_{2m})/2 + 2(f_1 + f_3 + \dots + f_{2m-1}) \\ + (f_2 + f_4 + \dots + f_{2m-2})), \end{aligned}$$

where $x_0 = a$, $x_{2m} = b$, $h = (b - a)/2m$, with truncation error $(b - a)^5 f^{(4)}/180(2m)^4$.

For the integral of a periodic analytic function over a single period the trapezoidal rule with equally spaced points gives the best result asymptotically, as the number of points tends to infinity.

Newton-Cotes formulas can also be obtained by integrating over the interval $[a, b]$ the interpolation polynomial for equally spaced points. The formulas mentioned so far are called **closed formulas** since they use values at the two endpoints of the interval. We can also use **open formulas**, which do not use values at the ends. For example, we have $\int_{x_0}^{x_4} f(x)dx = 4h(2f_1 - f_2 + 2f_3)/3$ with truncation error $14h^5 f^{(4)}/45$. Open formulas are useful for numerical solution of differential equations. There are also formulas that include values outside the interval, for example, $(-f_{-1} + 13f_0 + 13f_1 - f_2)h/24$ for $\int_{x_0}^{x_1} f(x)dx$. Applying these formulas to the m subintervals of $[a, b]$, we obtain a trapezoidal formula with

the correction term $h(f_1 - f_{-1})/24 + h(f_m - f_{m+1})/24 = h(\Delta f_0 + \Delta f_{-1})/24 - h(\Delta f_{m-1} + \Delta f_m)/24$, where Δf_i means $f_{i+1} - f_i$.

(2) Chebyshev Formulas. The Chebyshev formulas are a family of integration formulas in which all the weights W_i of $\sum_{i=1}^n W_i f(x_i)$ are equal, while the abscissas x_i are chosen so that the integral can be evaluated exactly when $f(x)$ is an arbitrary polynomial whose degree does not exceed n . When $\int_{-1}^1 f(x)dx = W \sum_{i=1}^n f(x_i)$, it is easy to see that $W = 2/n$ since the right-hand side must be equal to the left-hand side when $f(x) \equiv 1$. It is known that the abscissas x_i for $n \leq 7$ and $n = 9$ are real, while for $n = 8$ and $n \geq 10$ at least one of the abscissas becomes complex. It is easy to see that Chebyshev formulas are interpolatory.

(3) Gauss Formulas. In the Gauss formulas both the weights W_i and the abscissas x_i are chosen so that we obtain the accurate value of the integral when the integrand is any polynomial whose degree does not exceed $2n - 1$. If we put $\Pi(x) = \prod_{i=1}^n (x - x_i)$, an arbitrary polynomial of degree $2n - 1$ can be expressed in the form

$$f(x) = \sum_{k=1}^n \frac{\Pi(x)}{(x - x_k)\Pi'(x_k)} f_k + \Pi(x) \sum_{k=0}^{n-1} a_k x^k,$$

where $f(x_k) = f_k$ and the first term is the †Lagrange interpolation polynomial. By the assumption that the integral of $f(x)$ with weight $w(x)$ equals $\sum_{k=1}^n W_k f_k$, we obtain $W_k = \int_a^b (w(x)\Pi(x)/(x - x_k)\Pi'(x_k))dx$ and the relations

$$\int_a^b w(x)\Pi(x)x^k dx = 0, \quad k = 0, 1, \dots, n - 1.$$

Accordingly, the abscissas x_i are determined as the roots of the polynomial $\Pi(x)$ of degree n that is orthogonal to x^k ($k = 0, 1, \dots, n - 1$) with respect to the weight function $w(x)$.

The following are typical examples of integration formulas of Gaussian type.

(i) **Gauss integration formulas** (in the narrow sense). For $w(x) = 1$ and the interval $[-1, 1]$, $\Pi(x)$ is the †Legendre polynomial $P_n(x) = (1/2^n n!)d^n(x^2 - 1)^n/dx^n$. The error is $(n!)^4 2^{2n+1} f^{(2n)}/(2n + 1)((2n)!)^3$.

(ii) **Gauss-Laguerre formulas.** For the weight $w(x) = \exp(-x)$ and the interval $[0, \infty)$, $\Pi(x)$ is the † Laguerre polynomial $L_n(x) = (\exp x) d^n (x^n \exp(-x)) / dx^n$.

(iii) **Gauss-Hermite formulas.** For the weight function $w(x) = \exp(-x^2)$ and the interval $(-\infty, \infty)$, $\Pi(x)$ is the † Hermite polynomial $H_n(x) = (-1)^n \exp x^2 \cdot d^n \exp(-x^2) / dx^n$.

(iv) **Gauss-Chebyshev formulas.** For the weight function $w(x) = (1 - x^2)^{-1/2}$ and the interval $[-1, 1]$, we use the Chebyshev polynomial $T_n(x) = 2^{-(n-1)} \cos(n \arccos x)$. In this case, the W_i are all equal to π/n .

From the definition of W_k we see that the Gauss formulas are interpolatory.

(4) **Clenshaw-Curtis Formulas.** Although the Gauss formula is in general more accurate than the Newton-Cotes formula with the same number of points of interpolation, the points of interpolation for a Gauss formula of any order are distinct from those of any other order except the point zero, which appears in all formulas of odd order. The **Clenshaw-Curtis formulas** are interpolatory, with the points chosen so that the distribution of the points is similar to that of the Gauss formula and such that, in proceeding from a computation of order n to that of order $2n$, all the function values evaluated in the former computation be used in the latter. For $w(x) = 1$, the interval $[-1, 1]$, and even n , the points x_k and the weights W_k are given by

$$x_k = \cos \frac{k\pi}{n}, \quad k = 0, 1, \dots, n;$$

$$W_0 = W_n = \frac{1}{n^2 - 1};$$

$$W_s = W_{n-s} = \frac{4}{n} \sum_{j=0}^{n/2} \frac{1}{1 - 4j^2} \cos \frac{2\pi js}{n},$$

$$s = 1, 2, \dots, \frac{n}{2}.$$

Σ'' means that the first and the last terms in the sum are to be halved. There are some different types of Clenshaw-Curtis formulas depending on the selection of the points x_k [1].

B. Integration Formulas Based on Variable Transformation

If the integrand has some singularity at the endpoint, any interpolatory formula based on interpolation with a polynomial does not give a good result. In such cases **integration formulas based on variable transformation** are quite effective.

(1) **IMT Formula.** The † Euler-Maclaurin formula is given by

$$\int_a^b f(x) dx = h \left\{ \frac{1}{2} f(a) + \sum_{k=1}^{n-1} f(a + kh) + \frac{1}{2} f(b) \right\} \\ - \sum_{r=1}^m \frac{h^{2r} B_{2r}}{(2r)!} \{ f^{(2r-1)}(b) \\ - f^{(2r-1)}(a) \} + R_m,$$

$$h = \frac{b-a}{n},$$

$$R_m = \frac{h^{2m+1}}{(2m)!} \int_0^1 B_{2m}(t) \left\{ \sum_{k=0}^{n-1} f^{(2m)}(a + kh + ht) \right\} dt,$$

where the $B_n(t)$ are † Bernoulli polynomials of degree n and the B_n are † Bernoulli numbers ($B_0 = 1, B_1 = -1/2, B_2 = 1/6, B_3 = 0, B_4 = -1/30, \dots$). This formula suggests that if the higher derivatives of the integrand vanish at the both endpoints, the error of the trapezoidal rule with equally spaced points becomes very small. The **IMT formula** is based on the idea of transforming the variable x of $\int_0^1 f(x) dx$ in such a way that all the derivatives of the new integrand vanish at both endpoints by taking $x = \varphi(t)$, $\varphi(t) = K^{-1} \int_0^t \psi(\tau) d\tau$, $K = \int_0^1 \psi(\tau) d\tau$, $\psi(\tau) = \exp(-\tau^{-1} - (1-\tau)^{-1})$. Then the trapezoidal rule with $h = 1/n$ is applied to the transformed integral to obtain $(1/Kn) \sum_{j=1}^n \psi(j/n) f(\varphi(j/n))$ [1]. The asymptotic expression of the error for the IMT formula is proportional to $\exp(-C\sqrt{n})$ with a positive constant C .

(2) **Double Exponential Formula.** The trapezoidal rule with equally spaced points applied to the integral of an analytic function over $(-\infty, \infty)$ gives in general a result of high accuracy. The **double exponential formula** is based on the idea of transforming $\int_{-1}^1 f(x) dx$ to $\int_{-\infty}^{\infty} f(\varphi(t)) \varphi'(t) dt$ with $x = \varphi(t) = \tanh(\frac{1}{2} \pi \sinh t)$ and applying the trapezoidal rule with a mesh size h , which results in $h \sum_{n=-\infty}^{\infty} f(\varphi(nh)) \varphi'(nh)$ [4]. The name of the double exponential formula is attributed to the decay of $\varphi'(t)$ at $t \rightarrow \pm \infty$, which is approximately proportional to $\exp(-C \exp|t|)$ with a positive constant C . The transformations $x = \exp(\pi \sinh t)$ and $x = \sinh(\frac{1}{2} \pi \sinh t)$ give the double exponential formulas for the infinite integrals $\int_0^{\infty} f(x) dx$ and $\int_{-\infty}^{\infty} f(x) dx$, respectively. In the actual computation the infinite summation is truncated at appropriate upper and lower bounds. The asymptotic expression of the error for the double exponential formula in terms of the number N of the sampling points actually used is proportional to $\exp(-CN/\log N)$ with a positive constant C . The IMT formula and the double exponential

formula are robust against the singularities at the endpoints.

C. Automatic Integration

By an **automatic integration scheme** we mean a computer program for numerical integration of $\int_a^b f(x)dx$ in which the user gives the limits of integration a and b , a subroutine for computing $f(x)$, and an error tolerance ε . Then the program gives a value of the integral which is expected to be correct within the tolerance ε . Usually in an automatic integration scheme the mesh size is halved until the desired accuracy is attained. Automatic integration schemes are classified into two groups, nonadaptive schemes and adaptive schemes. In a **nonadaptive scheme**, the sequence of integration points is chosen according to some fixed rule independent of the shape of the integrand. Newton-Cotes formulas, Clenshaw-Curtis formulas, IMT formulas, and double exponential formulas can be used as base formulas for nonadaptive schemes.

From the historical point of view, **Romberg integration** should be mentioned; it is a kind of nonadaptive automatic integrator. Consider an integral $I = \int_a^b f(x)dx$. Divide the interval $[a, b]$ of integration into 2^k subintervals and apply the trapezoidal rule with the mesh size $h = (b-a)/2^k$, which we denote by $T_0^{(k)}$. Then, starting from the values obtained for $T_0^{(k)}$, $k=0, 1, \dots$, we compute the sequence

$$T_m^{(k)} = \frac{4^m T_{m-1}^{(k+1)} - T_{m-1}^{(k)}}{4^m - 1}, \quad m = 1, 2, \dots$$

From the † Euler-Maclaurin formula the asymptotic error for the trapezoidal rule $T_0^{(k)}$ can be expressed as $I - T_0^{(k)} = c_1 h^2 + c_2 h^4 + \dots + c_m h^{2m} + \dots$, where $c_m = \text{const} \times (f^{(2m-1)}(b) - f^{(2m-1)}(a))$ does not depend on h . If we compute $T_1^{(k)} = (4T_0^{(k+1)} - T_0^{(k)})/3$ using the values $T_0^{(k+1)}$ and $T_0^{(k)}$, then we see that the asymptotic error expression of $T_1^{(k)}$ becomes $I - T_1^{(k)} = c'_2 h^4 + c'_3 h^6 + \dots + c'_m h^{2m} + \dots$. Romberg integration is based on the idea of eliminating the term with h^{2m} in the expression of the error by successive application of $T_m^{(k)}$. This is an application of Richardson's extrapolation procedure.

In the **adaptive scheme**, the points are chosen in a way that depends on the shape of the integrand. The Newton-Cotes formula of order 8, for example, is used as the base formula for the adaptive scheme.

D. Approximate Multiple Integration

If a region is a product region, such as a rectangular parallelepiped, a **product rule** ob-

299 Ref.

Numerical Integration

tained by forming the product of 1-dimensional formulas is useful. For integrals over a cube or a sphere there are monomial rules which are exact for a certain family of monomials. These rules can be used for integrals of dimension lower than 5 or 6. For integrals of higher dimension only methods based on sampling make sense.

E. Numerical Differentiation

In order to find the numerical value of the derivative $f'_p = f'(x_p)$ at a point x_p from the tabulated values $f_k = f(x_k)$, we usually use the derivative of the † Lagrange interpolation formula. This gives

$$f'_p = \sum_{k \neq p} \frac{\Pi'(x_p)}{(x_p - x_k)\Pi'(x_k)} f_p + \sum_{k \neq p} \frac{1}{x_p - x_k} f_k,$$

where $\Pi(x) = \prod_{k=1}^n (x - x_k)$.

When we compute the derivative of a function which can be evaluated at any point in a given interval, the approximation

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h}$$

is useful; similarly, we can use

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

It must be noted that, as h tends to zero, the difference $f(x+h) - f(x-h)$ comes to contain fewer significant digits, so that it is meaningless to carry out $\{f(x+h) - f(x-h)\}/2h$ beyond a reasonable value of h .

References

- [1] P. J. Davis and P. Rabinowitz, *Methods of numerical integration*, Academic Press, 1975.
- [2] P. J. Davis and I. Polonsky, *Numerical interpolation, differentiation and integration*, *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun (eds.), Nat. Bur. Stand. Appl. Math. Ser., 55 (1964), 875-924.
- [3] V. I. Krylov, *Approximate calculation of integrals*, translated by A. H. Stroud, Macmillan, 1962.
- [4] H. Takahasi and M. Mori, *Double exponential formulas for numerical integration*, *Publ. Res. Inst. Math. Sci.*, 9 (1974), 721-741.
- [5] J. R. Rice, *A meta-algorithm for adaptive quadrature*, *J. ACM*, 22 (1975), 61-82.
- [6] A. H. Stroud and D. H. Secrest, *Gaussian quadrature formulas*, Prentice-Hall, 1966.
- [7] A. H. Stroud, *Approximate calculation of multiple integrals*, Prentice-Hall, 1971.

300 (XV.1) Numerical Methods

In the earlier history of mathematics, the development of methods of numerical calculation was one of the main purposes of research. Until the beginning of this century, logarithmic calculation played a central role in numerical calculation, and the main topic of this field was to make tables of values of functions. The digital electronic computer (\rightarrow 75 Computers), which made its first appearance in the 1940s and has been developing at an exponential rate, has caused drastic changes in numerical technique. Problem solving by numerical methods has now become one of the fundamental means of research in the physical sciences and engineering, and also in the social sciences and humanities. In this article we give examples of the changes in numerical methods brought about by the availability of digital computers and portable calculators.

Computers may be effectively utilized for calculating individual values of functions. This has led to the reexamination and, in some cases, modification of approximate formulas for evaluating functions (\rightarrow 142 Evaluation of Functions). For familiar functions, such as \dagger logarithmic, \dagger exponential, and \dagger trigonometric functions, tables have been almost completely replaced by function keys on electronic calculators. Microprogramming algorithms for obtaining values of these functions have also been devised [1]. A complex function-theoretic error-estimation method for use with numerical integration formulas is given in [2]. In this method, graphical outputs of the computer are utilized. For problems in which the existence and uniqueness of solutions have been established, as for algebraic equations and ordinary differential equations, fairly good numerical calculation methods and their error estimates have been established (\rightarrow 301 Numerical Solution of Algebraic Equations; 303 Numerical Solution of Ordinary Differential Equations; [3,4]). A method for estimating arithmetic \dagger accumulation errors for operations involving finite numbers of digits has been systematized (\rightarrow 138 Error Analysis; [5]).

It is not unusual nowadays for linear equations with 10,000 or more unknowns to be solved. As new computing systems, such as the virtual memory system and the vector operation system, come into existence, new algorithms are examined; a numerical method that is optimal for today's technology may well be suboptimal for the next generation of computers. General-purpose program packages of linear problems, including eigenvalue problems, have been developed (\rightarrow 298

Numerical Computation of Eigenvalues; [6]).

Partial differential equations seem to have become more familiar because of the visualization of their numerical solutions in graphical computer outputs. In [7], which appeared much earlier than computers, partial difference equations for the fundamental linear problems in mathematical physics were discussed with a suggested variational treatment (\rightarrow 304 Numerical Solution of Partial Differential Equations; [8]). The \dagger finite element method, which started as a calculation technique in structural mechanics and is based on the \dagger calculus of variations, is widely accepted as an efficient approximation method for partial differential equations [9, 10]. The term \dagger simulation is used often to describe procedures in which partial differential equations describing time-dependent phenomena are discretized; the resulting difference system can be solved for long periods of time [11].

Numerical analysis has heretofore long been considered to be the only numerical method (for error analysis in particular), and has been carried out mainly by means of the discretization of equations. Nowadays, however, **mathematical modeling**, taking into account both the phenomena to be described and the capabilities of the computers to be used, has become an important numerical method.

References

- [1] J. S. Walther, A unified algorithm for elementary functions, Conference Proceedings, Spring Joint Computer Conference, AFIPS 38 (1971), AFIPS Press, 1971, 379–385.
- [2] H. Takahashi and M. Mori, Error estimation in the numerical integration of analytic functions, Rep. Computing Centre, Univ. of Tokyo, 3 (1970), 41–108.
- [3] P. Henrici, Elements of numerical analysis, Wiley, 1974.
- [4] P. Henrici, Discrete variable methods in ordinary differential equations, Wiley, second edition, 1968.
- [5] J. H. Wilkinson, Rounding errors in algebraic processes, HMSO, 1963.
- [6] J. H. Wilkinson and C. Reinsch (eds.), Handbook for automatic computation 2, pt. II, Linear algebra, Springer, 1971.
- [7] R. Courant, K. Frienrichs, and H. Lewy, On the partial difference equations of mathematical physics, IBM J. Res. Develop., 11 (1967), 215–234. (Original in German, 1928.)
- [8] F. John, Lectures on advanced numerical analysis, Gordon & Breach, 1967.
- [9] G. Strang and G. J. Fix, An analysis of the finite element method, Prentice-Hall, 1973.

[10] P. G. Ciarlet, The finite element method for elliptic problems, North-Holland, 1978.

[11] R. J. Roache, Computational fluid dynamics, Hermosa, 1972.

301 (XV.5) Numerical Solution of Algebraic Equations

A. General Remarks

Methods for finding roots of an equation $f(x)=0$ (where $f(x)$ is not necessarily a polynomial, but is assumed to be a function with some regularity) by numerical calculation can in general be divided into the following two types: (i) The first has as its goal the finding of good approximate values of the roots; examples are the Bernoulli method (Section J) and the Graeffe method (Section N). (ii) The second improves the accuracy of estimates of the roots; an example is the Newton-Raphson method (Section D). The methods belonging to (i) and (ii) can be used separately. However, in (i) convergence of the approximations may be excessively slow when a pair of nearly equal roots is present, and the size of numbers involved in the calculation may grow exponentially as we proceed through the iterations. In (ii) convergence of the approximate values is not assured unless the initial approximate value is suitably chosen. Accordingly, when an electronic computer is utilized, it is advisable to combine both types of method. Because of the development of computers, solutions with global convergence have become important [1, 2, 15, 17, 21].

B. Successive Substitutions

When an equation $f(x)=0$ is transformed into $x=F(x)$ and the roots are obtained by iterative calculation of $x_{i+1}=F(x_i)$ ($i=0, 1, 2, \dots$), sufficient conditions for its convergence are as follows: Let one of the roots of the equation be α . Then $x_{i+1}-\alpha=F(x_i)-\alpha$, and therefore $(x_{i+1}-\alpha)/(x_i-\alpha)=F'(\xi)$ ($x_i < \xi < \alpha$ or $x_i > \xi > \alpha$). Accordingly, α converges monotonically if $0 < F'(\xi) < 1$, while it converges with oscillation if $0 > F'(\xi) > -1$. If $|F'(\xi)| > 1$ and $F^{-1}(x)$ denotes the inverse function of $F(x)$, the iteration $x_{i+1}=F^{-1}(x_i)$ converges. To define the speed of convergence by iterative processes the following notion of order is utilized: When $\lim_{i \rightarrow \infty} x_i = \alpha$, the speed of convergence of the sequence $\{x_i\}$ is said to be of the k th order if $\lim_{i \rightarrow \infty} (x_{i+1} - \alpha)/(x_i - \alpha)^k = c \neq 0$. Necessary and sufficient

conditions for the speed of convergence of $\{x_i\}$ to be of the k th order are $\alpha = F(\alpha)$, $F'(\alpha) = F''(\alpha) = \dots = F^{(k-1)}(\alpha) = 0$, $F^{(k)}(\alpha) \neq 0$.

The main iterative processes used in numerical calculation are described in the following sections.

C. Regula Falsi

Regula falsi (or the **method of false position**) is a process of obtaining the real root α of an equation $f(x)=0$ by approaching the root from both sides. The calculation procedure is as follows: Assume that $f(x_p) > 0$ and $f(x_q) < 0$, where x_p and x_q are approximate values of α such that α lies between them. A new approximate value \bar{x} is then obtained from $\bar{x} = (x_q f(x_p) - x_p f(x_q)) / (f(x_p) - f(x_q))$. If $f(\bar{x}) > 0$, then $\bar{x} \rightarrow x_p$, and if $f(\bar{x}) < 0$, then $\bar{x} \rightarrow x_q$, and the procedure is repeated (the symbol \rightarrow means replacement). The conditions and speed of convergence of regula falsi are as follows: Let $F(x) = (x_q f(x) - x f(x_q)) / (f(x) - f(x_q))$; then $F'(\alpha) = (f(x_q) + (\alpha - x_q)f''(\alpha)) / f(x_q)$. If f' and f'' exist and are continuous near α , then $f(x_q) = f(\alpha) + (x_q - \alpha)f'(\alpha) + (1/2)(x_q - \alpha)^2 f''(\xi)$, where $\alpha > \xi > x_q$ or $\alpha < \xi < x_q$. Accordingly, assume that $F'(\alpha) = (1/2)(x_q - \alpha)^2 f''(\xi) / f(x_q) \neq 0$. Then $|F'(x)| < 1$ can be satisfied if x and x_q are near α . Therefore, if the initial value is appropriate, the convergence is of the first order. In regula falsi, only calculation of $f(x)$ is necessary, while that of $f'(x)$ is unnecessary. Furthermore, if $f(x)=0$ has nearby real roots, i.e., α, β, \dots close to each other and $f'(x)$ small near the roots, there can be no mistake such as the neighboring root β being obtained in the process of finding the root α by this method. This method belongs to the inverse linear interpolation method. This kind of inverse interpolation includes Muller's method [3], which uses † Lagrange's interpolation formula, the Torii-Miyakoda method [4], which uses † Hermite's interpolation formula, and Whittaker's method [5], which uses † Stirling's interpolation formula. These methods converge more rapidly but have more complicated formulas than inverse linear interpolation. The **Sturm method** (in which the interval where the roots exist is narrowed by † Sturm's theorem) and the **Horner method** (which obtains the decimal part digit by digit), both used to obtain the real roots of a high-order algebraic equation $f(x)=0$, are also of this type.

D. The Newton-Raphson Method

In obtaining the real roots of an equation $f(x)=0$, the **Newton-Raphson method** (or the **New-**

ton iterative process), which converges rapidly, is used when $f'(x) \neq 0$ is computable. Let x_i be a sufficiently close i th approximation of the root α ; then $x_{i+1} = x_i - f(x_i)/f'(x_i)$ is closer than x_i to the true solution. The process is repeated until $|x_{i+1} - x_i|$ is small enough. Conditions and the speed of convergence are as follows [6, 7]: Let $F(x) = x - f(x)/f'(x)$; then $F'(x) = f(x)f''(x)/(f'(x))^2$ and $F'(\alpha) = 0$. Accordingly, $F''(\alpha) \neq 0$ if $f'(\alpha) \neq 0$ and $f''(\alpha) \neq 0$, and the convergence is of the second order if it is possible to determine an appropriate approximate value x_i that is close to α and satisfies $|F'(x_i)| < 1$. In particular, if $f(x_0)f'(x_0) \neq 0$, $h_0 = -f(x_0)/f'(x_0)$, $|f''(x)| \leq M$, and $|f'(x_0)| \geq 2|h_0|M$, then every Newton approximation x_i starting with x_0 is contained in the interval $I = [x_1 - |h_0|, x_1 + |h_0|]$, the equation has only one root α in I , and $x_i \rightarrow \alpha$. Besides,

$$|\alpha - x_{i+1}| \leq M|x_i - x_{i-1}|^2/2|f'(x_i)|, \quad i = 1, 2, \dots$$

With regard to convergence and evaluation of the error, including roundoff error in practical computations, M. Urabe's studies should be consulted [8, 9]. In general, the convergence is of the third order in the Newton-Raphson method, which uses not only $f'(x)$ but also $f''(x)$ [10].

When $f'(x) = 0$, we must assume $f''(x) \neq 0$ and use the value $f''(x)$. With regard to this case, the study by S. Hitotumatu should be consulted [11]. In addition, W. Kizner (*SIAM J. Appl. Math.*, 12 (1964)) reported an iterative process in which the convergence is of the fifth order without using any derived function higher in order than $f''(x)$. The essential part of his method lies in a numerical integration of the integral part by the †Runge-Kutta formula, based on the fact that if x_1 is the first approximate value of a root \bar{x} of $f(x) = 0$, then

$$\bar{x} = \int_{f(x_1)}^0 \frac{dx}{df} df + x_1.$$

There may be several iterative processes for solving $f(x) = 0$, even if the order k of the convergence is fixed. For example, in obtaining the positive root of $f(x) = x^2 - a = 0$ ($a > 0$), i.e., the square root $\alpha = a^{1/2}$, both iterative processes $x_{i+1} = (x_i + a/x_i)/2$ and $x_{i+1} = 2x_i^3/(3x_i^2 - a)$ give convergence of the second order. In obtaining the real root of $f(x) = x^3 - a = 0$, i.e., the cube root $\alpha = a^{1/3}$, $x_{i+1} = x_i + (a/x_i^2 - x_i)/3$ converges of the second order, while $x_{i+1} = x_i/2 + (a + a/2)/(2x_i^2 + a/x_i)$ converges of the third order.

The Newton-Raphson method is applicable also to holomorphic functions of complex variables.

Take simultaneous equations of two variables, $f(x, y) = 0$ and $g(x, y) = 0$. If it is possible to transform the equations into $x = \varphi(x, y)$ and

$y = \psi(x, y)$, and if $|\partial\varphi/\partial x| + |\partial\psi/\partial x| < 1$, $|\partial\varphi/\partial y| + |\partial\psi/\partial y| < 1$, then the iterative processes $x_{i+1} = \varphi(x_i, y_i)$ and $y_{i+1} = \psi(x_i, y_i)$ are applicable. In the Newton-Raphson method, let the corrections to the i th approximate values x_i and y_i be Δx_i and Δy_i , and solve

$$f_x(x_i, y_i)\Delta x_i + f_y(x_i, y_i)\Delta y_i = -f(x_i, y_i),$$

$$g_x(x_i, y_i)\Delta x_i + g_y(x_i, y_i)\Delta y_i = -g(x_i, y_i).$$

Then we can take $x_{i+1} = x_i + \Delta x_i$ and $y_{i+1} = y_i + \Delta y_i$ as the next approximate values. This process is repeated until both $|\Delta x_i|$ and $|\Delta y_i|$ are sufficiently small. More generally, the Newton-Raphson method for solving a system of n nonlinear equations $f_i(x_1, \dots, x_n) = 0$, $i = 1, 2, \dots, n$, is defined by

$$x^{(k+1)} = x^{(k)} - J(x^{(k)})^{-1}F(x^{(k)}), \quad k = 0, 1, 2, \dots,$$

where $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})^t$, $F(x) = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))^t$, $J(x) = (\partial f_i(x_1, \dots, x_n)/\partial x_j)$ (the †Jacobian matrix of F), and where $x^{(0)}$ is chosen appropriately [23].

E. The Bairstow Method

Corresponding to a pair of complex roots $\alpha \pm i\beta$ of an algebraic equation $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_n = 0$ with real coefficients, there is a real quadratic factor $x^2 + p^*x + q^*$. The **Bairstow method** (or **Hitchcock method**) obtains the coefficients p^* and q^* of this quadratic factor.

To do this, first choose an appropriate quadratic factor $x^2 + px + q$ as a candidate. Then through synthetic division by the quadratic factor, b_j and c_j are computed by means of the formulas

$$b_j = a_j - pb_{j-1} - qb_{j-2}, \quad j = 0, 1, \dots, n,$$

and

$$c_j = b_j - pc_{j-1} - qc_{j-2}, \quad j = 0, 1, \dots, n-1,$$

where

$$b_{-1} = b_{-2} = c_{-1} = c_{-2} = 0.$$

By solving the simultaneous equations

$$c_{n-2}\Delta p + c_{n-3}\Delta q = b_{n-1},$$

$$\bar{c}_{n-1}\Delta p + c_{n-2}\Delta q = b_n,$$

where

$$\bar{c}_{n-1} = c_{n-1} - b_{n-1},$$

the quantities Δp and Δq are obtained. Then taking $\bar{p} = p + \Delta p$, $\bar{q} = q + \Delta q$, we have $x^2 + \bar{p}x + \bar{q}$ as a new candidate. This operation is repeated until Δp and Δq are sufficiently small. Since this method corresponds to the Newton-Raphson method in the case of two variables, and, accordingly, the convergence is of the

second order, a choice of suitable initial values of p and q leads to rapid convergence. The key to this method lies in choosing p and q so that $R_1(p, q) = 0$ and $R_2(p, q) = 0$, where R_1 and R_2 are such that $R_1x + R_2$ is the remainder of $f(x)$ divided by the trial quadratic factor $x^2 + px + q$. This method was generalized by A. A. Grau (*SIAM J. Appl. Math.*, 11 (1963)). Namely, when $f(x) = (x^2 + px + q)g(x) + r(x)$ with $r(x) = r_1x^{k+1} + r_2x^k$, the functions r_1 and r_2 can be used instead of R_1 and R_2 . The process corresponds to the Bairstow method when $k = 0$, and to the McAuley method (*SIAM J. Appl. Math.*, 10 (1962)) when $k = n - 2$.

F. The Durand-Kerner Method

The **Durand-Kerner (DK) method** [12] for solving an algebraic equation $f(z) = z^n + a_1z^{n-1} + \dots + a_n = 0$ ($a_n \neq 0$) with complex coefficients is an iterative method defined by

$$z_i^{(k+1)} = z_i^{(k)} - \frac{f(z_i^{(k)})}{\prod_{j=1, j \neq i}^n (z_i^{(k)} - z_j^{(k)})}, \quad i = 1, 2, \dots, n, \quad k = 0, 1, 2, \dots \quad (1)$$

A feature of this method is that it can determine simultaneously all the roots of $f(z) = 0$. Let $\varphi_m(z_1, \dots, z_n)$, $m = 1, 2, \dots, n$, be the m th elementary symmetric functions with respect to z_1, \dots, z_n :

$$\varphi_m(z_1, \dots, z_n) = \sum_{i_1 < i_2 < \dots < i_m} z_{i_1} z_{i_2} \dots z_{i_m}.$$

Set $f_m(z) = (-1)^m \varphi_m(z_1, \dots, z_n) - a_m$. Then $\alpha = (\alpha_1, \dots, \alpha_n)^t$, a set of the roots of $f(z) = 0$, is a solution of a system of the n equations $f_m(z) = 0$, $m = 1, 2, \dots, n$. I. O. Kerner [12] showed that the DK method is the Newton-Raphson method applied to the system of nonlinear equations $f_m(z) = 0$, $m = 1, 2, \dots, n$. Therefore the speed of convergence of (1) is of the second order, provided that it converges. The initial values $z_1^{(0)}, \dots, z_n^{(0)}$ for the DK method are usually chosen as follows: Let $g(z) = f(z - a_1/n) = z^n + c_2z^{n-2} + \dots + c_n$ and $h(z) = z^n - |c_2|z^{n-2} - \dots - |c_n|$. If $(c_2, \dots, c_n) \neq (0, \dots, 0)$ ($n \geq 2$), then it can be shown that $h(z) = 0$ has exactly one positive root r , and all the roots of $f(z) = 0$ lie in the disk $|z + a_1/n| \leq r$. Now, with a positive $r_0 \geq r$ and $\theta = \pi/(2n)$, put

$$z_i^{(0)} = -\frac{a_1}{n} + r_0 \exp \left[\left(\frac{2(i-1)\pi}{n} + \theta \right) \sqrt{-1} \right], \quad i = 1, 2, \dots, n. \quad (2)$$

Such a choice was proposed by O. Aberth [13]. Hence the process (1) together with (2) can be called the **Durand-Kerner-Aberth (DKA) method**. It is shown for the DKA

method that, if r_0 is large enough, then

$$z_i^{(k)} + \frac{a_1}{n} = \left(1 - \frac{1}{n} \right)^k \left(z_i^{(0)} + \frac{a_1}{n} \right), \quad i = 1, 2, \dots, n,$$

hold nearly for the first several steps. Thus the DKA method has a certain kind of global convergence property that renders it one of the most powerful methods for solving algebraic equations.

A variant of the Durand-Kerner method is the Ehrlich-Aberth method [13, 14]; it is defined for $i = 1, 2, \dots, n$, $k = 0, 1, 2, \dots$ by

$$z_i^{(k+1)} = z_i^{(k)} - \frac{f(z_i^{(k)})}{f'(z_i^{(k)}) - f(z_i^{(k)}) \sum_{j=1, j \neq i}^n 1/(z_i^{(k)} - z_j^{(k)})}.$$

The speed of convergence of this method is of the third order. Further variants of the Durand-Kerner and the Ehrlich-Aberth methods have been proposed by M. Iri et al. [15] and A. W. M. Nourain [16].

G. The Dejon-Nickel Method

Let $f(z)$ be a polynomial of degree n with complex coefficients, and take z_0 such that $f(z_0) \neq 0$. Then we can write

$$f(z_0 + h) = f(z_0) \{ 1 + b_1 h^i + b_{i+1} h^{i+1} + \dots + b_n h^n \} \quad (b_i \neq 0) \\ = f(z_0) \{ 1 + b_i h^i (1 + \theta) \}, \quad \theta = o(h).$$

Now, choose h sufficiently small so that $|b_i h^i| < 1$, $|\theta| < 1$, and $\arg(b_i h^i) = \pi$. Then $|1 + b_i h^i| = 1 - |b_i h^i|$ and $|b_i h^i \theta| < |b_i h^i|$ so that $|f(z_0 + h)| \leq |f(z_0)| \{ |1 + b_i h^i| + |b_i h^i \theta| \} < |f(z_0)|$. Hence, if z_0 is chosen so that $\min |f(z)| = |f(z_0)|$, then we must have $f(z_0) = 0$. This is the outline of Cauchy's existence proof for the roots of algebraic equations. The **Dejon-Nickel method** [17] is a method in which h is chosen as follows:

$$h = (-1/b_k)^{1/k},$$

where

$$\min \{ |1/b_j|^{1/j} \mid b_j \neq 0, i \leq j \leq n \} = |1/b_k|^{1/k} \quad (=r, \text{ say}).$$

If several such k exist, take the smallest one.

The branch of the multivalued function $(-1/b_k)^{1/k}$ is chosen such that $z^{1/k}$ is positive for z positive. If $|f(z_0 + h)| \leq (1 - \varepsilon)|f(z_0)|$ with a preassigned constant ε such that $0 < \varepsilon < 1$, then put $z_1 = z_0 + h$. If the inequality does not hold, then for each integer $m \geq 1$ choose the smallest integer $l = l(m)$ such that $\max \{ |b_j| (r/2^m)^j \mid i \leq j \leq n \} = |b_l| (r/2^m)^l$, and put $\tilde{z}_m = z_0 + (r/2^m) (-|b_l|/b_l)^{1/l}$. Find the smallest integer $m \geq 1$ such that $|f(z_0)| \leq (1 - 2^{-1} (r/2^m)^l |b_l|) |f(z_0)|$, and put

$z_1 = \tilde{z}_m$. By continuing this process, a sequence $\{z_j\}$ is constructed such that $|f(z_0)| > |f(z_1)| > |f(z_2)| > \dots$. It converges toward some root of the equation $f(z) = 0$. S. Hirano has proposed a similar method.

H. Methods for Finding Good First Estimates of Roots

Some of the principal methods for obtaining good first approximate values of the roots are given in the following sections I–N.

I. Matrix Methods

The problem of solving an algebraic equation $f(z) = z^n + a_1 z^{n-1} + \dots + a_n = 0$ with complex coefficients is equivalent to that of finding the eigenvalues of the **companion matrix**

$$A = \begin{bmatrix} -a_1 & -a_2 & -a_3 & \dots & -a_{n-1} & -a_n \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Therefore numerical methods for solving the eigenvalue problems for nonsymmetric matrices are applicable (\rightarrow 298 Numerical Computation of Eigenvalues). However, it should be remarked that this might be inefficient because most of the elements of the matrix A are zero [18].

J. The Bernoulli Method

In the **Bernoulli method** the iterative formulas $S_1 = -a_1$, $S_k = -(a_1 S_{k-1} + a_2 S_{k-2} + \dots + a_{k-1} S_1 + k a_k)$ ($k = 2, 3, \dots, n$), $S_k = -a_1 S_{k-1} - a_2 S_{k-2} - \dots - a_n S_{k-n}$ ($k = n+1, n+2, \dots$) are repeatedly applied to an algebraic equation $f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_n = 0$. When the roots of $f(x) = 0$ are $\alpha_1, \alpha_2, \dots, \alpha_n$ ($|\alpha_1| \geq |\alpha_2| \geq \dots \geq |\alpha_n|$), $S_k/S_{k-1} \rightarrow \alpha_1$ if α_1 is a real and simple root and $|\alpha_1| > |\alpha_2|$. When α_1, α_2 are complex roots, we put $\alpha_1 = R e^{i\theta}$, $\alpha_2 = R e^{-i\theta}$. If $|\alpha_3| < R$, then we have

$$\frac{S_k^2 - S_{k+1} S_{k-1}}{S_{k-1}^2 - S_k S_{k-2}} \rightarrow R^2,$$

$$\frac{S_k S_{k-1} - S_{k+1} S_{k-2}}{S_{k-1}^2 - S_k S_{k-2}} \rightarrow 2R \cos \theta,$$

and hence the two roots of $x^2 - (2R \cos \theta)x + R^2 = 0$ are α_1 and α_2 . S_k is the sum of the k th powers of the n roots of the equation, $\alpha_1, \alpha_2, \dots, \alpha_n$. C. Lanczos used $x f'(x)/f(x) = n + S_1/x + S_2/x^2 + \dots + S_k/x^k + \dots$ to compute S_k [10, pp. 26–30]. Let the eigenvalues of the companion matrix A be $\lambda_1, \lambda_2, \dots, \lambda_n$ ($|\lambda_1| \geq |\lambda_2| \geq$

$\dots \geq |\lambda_n|$). Then the Bernoulli method is the same as the \dagger power method for obtaining the maximum characteristic value λ_1 (\rightarrow 298 Numerical Computation of Eigenvalues C).

K. The Lehmer Method

The **Lehmer method** [9] is a general method for finding the n roots of an algebraic equation with complex coefficients

$$f(z) = a_0 + a_1 z + a_2 z^2 + \dots + a_n z^n = 0$$

in the complex plane by repeating the following procedure: First, draw a circle whose center is at the origin and whose radius is $R = r 2^{-\theta}$, where r is an arbitrary given number and θ an arbitrary given integer. Then, utilizing the Lehmer theorem, observe whether a root α of $f(z) = 0$ lies inside the circle. If there is no such root, replace R with $2R$, whereas if such a root exists, replace R with $R/2$. Continue this process until an R for which a root α exists within the annulus $R < |z| < 2R$ is obtained.

Second, draw circles with the centers $\beta_k = (5R/3) \exp(i 2\pi k/8)$ ($k = 0, 1, \dots, 7$) and common radius $\mu = (5R/3)/2$ and cover the annulus obtained by the first step. Then find out which of the circles has the root α in its interior. If α is in the interior of the circle for $k = j$, the origin is shifted to β_j and the operation is started again from the first step. If R_1 satisfying $R_1 < |z - \beta_j| < 2R_1$ is obtained by the first step, we have $R_1 \leq (5/12)R$. Therefore, when the first step is repeated N times, the root α is confined in a small circle whose radius is smaller than $2(5/12)^N R$. Then the center β of this small circle gives a good approximate value of α .

L. The Downhill Method

The downhill method is a method for obtaining the extreme values of a function of many variables. It is also applicable in obtaining approximate solutions of a system of equations. Let us consider the **downhill method** in the case of two variables. The problem of obtaining the real roots of simultaneous equations $f(x, y) = 0$ and $g(x, y) = 0$ can be reduced to the problem of obtaining the coordinates (α, β) which give the extreme value 0 of $\Phi(x, y)$, where $\Phi(x, y) = f^2 + g^2$. The values of $\Phi(x, y)$ are calculated at 3^2 points obtained by the combination of $x = x_1, x_1 \pm h; y = y_1, y_1 \pm h$, where (x_1, y_1) are arbitrary approximate values of (α, β) constituting the centers of those sets of points, and h is the given step size. Utilizing the values of the function at the 3^2 points,

$\Phi(x, y)$ is approximated by a quadratic surface

$$b_0 + b_1x + b_2y + b_{11}(3x^2 - 2) + b_{22}(3y^2 - 2) + b_{12}xy.$$

The values of $b_0, b_1, b_2, b_{11}, b_{22}, b_{12}$ are calculated by the †method of least squares, and the center (x_1^*, y_1^*) of the approximate quadratic surface is obtained. Then the first approximations x_1 and y_1 are replaced by the second approximations $x_1 + x_1^*$ and $y_1 + y_1^*$. This process is repeated while the step size h is made appropriately smaller. This method is an improvement on the method of successive experimental planning used by G. E. P. Box and K. B. Wilson (1954) to obtain optimum conditions in the exploration of response curved surfaces.

By using another minimization technique, J. A. Grant and G. D. Hitchins [21] gave an ever-convergent algorithm for determining initial approximations to the roots of algebraic equations with real coefficients. The practical implementation of this method is given in J. A. Grant and G. D. Hitchins (*Comput. J.*, 18 (1975)).

M. Continuation Methods

Let $F(x) = (f_1(x), \dots, f_n(x))^t = 0$ ($x = (x_1, \dots, x_n)^t$) be a system of n equations. Suppose that no reasonable approximation for a solution exists. Then, take arbitrary $x^{(0)}$ and define a one-parameter family of equations

$$H(x, t) \equiv F(x) - (1-t)F(x^{(0)}) = 0, \quad 0 \leq t \leq 1. \quad (3)$$

Suppose that for each t the equation has a solution $x(t)$ which depends continuously on t . Observe then that $x(0) = x^{(0)}$ and $x(1)$ is a solution for the equation $F(x) = 0$. Partition the interval $[0, 1]$ by the points $0 = t_0 < t_1 < \dots < t_N = 1$. First, solve $H(x, t_1) = 0$ by some iterative method using $x^{(0)}$ as a first approximation. Let $x^{(1)}$ be a solution thus obtained. Next, solve $H(x, t_2) = 0$ by some iterative method using $x^{(1)}$ as a first approximation, and so on. Finally, solve $H(x, t_N) = 0$ by some iterative method which uses $x^{(N-1)}$ as a first approximation. Then $x^{(N)}$, a solution thus obtained, can be used as a first approximation in an iterative method applied to the equation $F(x) = 0$. This method is called a **continuation method** [22–24].

As another approach, differentiate (3) with respect to t . Then $J(x(t))x'(t) + F(x^{(0)}) = 0$, where $J(x(t))$ is the Jacobian matrix of F , evaluated at $x = x(t)$. Hence $x(t)$ is the solution of the system of ordinary differential equations $x'(t) = -J(x(t))^{-1}F(x^{(0)})$, $0 \leq t \leq 1$,

subject to the initial condition $x(0) = x^{(0)}$, provided that $J(x(t))$ is nonsingular. Therefore numerical solution at $t = 1$ gives a good approximation for a solution of $F(x) = 0$. This is called **Daivenko's method of differentiation with respect to a parameter**. These methods can be used to find initial approximations of the Durand-Kerner and the Ehrlich-Aberth methods. They are also applicable to a single algebraic equation $f(x) = 0$, which is the special case $n = 1$.

N. Other Methods

To obtain the first approximate values of the roots of an algebraic equation, the Graeffe method has been used, in which the roots of the equation are separated by successively forming an equation whose roots are the squares of the roots of the preceding equation. In computer calculations, however, the other methods described in previous sections are more convenient than the Graeffe method.

The **Lanczos method** [10] is a method in which $g(x)$, an approximate function of $f(x)$, is obtained by the process of approximating functions, and the roots of $g(x) = 0$ are taken as approximate values of the roots of $f(x) = 0$. The Garside-Jarratt-Mack method [25] is a modification of the Lanczos method and approximates $f(x)/f'(x)$ by a rational function $g(x) = (x-a)/(b+cx)$.

O. Error Bounds for Computed Solutions

Let z_1, \dots, z_n be computed solutions of an algebraic equation $f(z) = 0$ which were obtained by some method. If z_1, \dots, z_n are distinct, then the following result due to B. T. Smith [26] is quite useful for estimating the errors of z_i : Let

$$\Gamma_i: |z - z_i| \leq \gamma_i \equiv \frac{n|f(z_i)|}{\prod_{j=1, j \neq i}^n |z_i - z_j|}.$$

Then the union of the disks Γ_i contains all the roots $\alpha_1, \dots, \alpha_n$ of $f(z) = 0$. Any connected component of $\bigcup_{i=1}^n \Gamma_i$, which consists of just m disks Γ_i , contains exactly m roots of $f(z) = 0$. Hence if $\Gamma_i \cap \Gamma_j = \emptyset$ for every $j \neq i$, then $|\alpha_i - z_i| \leq \gamma_i$. Smith obtained a more general result for the case where z_1, \dots, z_n are not necessarily distinct.

References

- [1] J. F. Traub, A class of global convergent iteration functions for solution of polynomial equations, *Math. Comp.*, 20 (1966), 113–138.
- [2] B. Dejon and P. Henrici (eds.), *Construc-*

tive aspects of the fundamental theorem of algebra, Wiley-Interscience, 1969.

- [3] D. E. Muller, A method for solving algebraic equations using an automatic computer, *Math. Tables Aids Comput.*, 10 (1956), 208–215.
- [4] T. Torii and T. Miyakoda, A practical root-finding algorithm based on the cubic Hermite interpolation, *Information Processing in Japan*, 13 (1973), 143–148.
- [5] E. T. Whittaker and G. Robinson, *Calculus of observations*, Van Nostrand, 1924.
- [6] L. B. Rall, *Computational solution of nonlinear operator equations*, Krieger, 1979.
- [7] A. M. Ostrowski, *Solution of equations in Euclidean and Banach spaces*, Academic Press, 1973.
- [8] M. Urabe, Error estimation in numerical solution of equations by iteration process, *J. Sci. Hiroshima Univ., ser. A.*, 26 (1962), 77–91.
- [9] M. Urabe, Component-wise error analysis of iterative methods practiced on a floating-point system, *Mem. Fac. Sci. Kyushu Univ.*, 27 (1973), 23–64.
- [10] C. Lanczos, *Applied analysis*, Prentice-Hall, 1956.
- [11] S. Hitotumatu, A method of successive approximation based on the expansion of second order, *Math. Japonicae*, 7 (1962), 31–50.
- [12] I. O. Kerner, Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen, *Numer. Math.*, 8 (1966), 290–294.
- [13] O. Aberth, Iteration methods for finding all zeros of a polynomial simultaneously, *Math. Comp.*, 27 (1973), 339–344.
- [14] L. W. Ehrlich, A modified Newton method for polynomials, *Comm. ACM*, 10 (1967), 107–108.
- [15] M. Iri, H. Yamashita, T. Terano, and H. Ono, An algebraic-equation solver with global convergence property, *Publ. Res. Inst. Math. Sci.*, 339 (1978), 43–69.
- [16] A. W. M. Nourain, An improvement on two iteration methods for simultaneous determination of the zeros of a polynomial, *Int. J. Computer Math.*, 6 (1977), 241–252.
- [17] B. Dejon and K. Nickel, A never failing, fast convergent root-finding algorithm, *Constructive Aspects of the Fundamental Theorem of Algebra*, B. Dejon and P. Henrici (eds.), Wiley-Interscience, 1969, 1–35.
- [18] D. M. Young and R. T. Gregory, *A survey of numerical mathematics I*, Addison-Wesley, 1972.
- [19] D. H. Lehmer, A machine method for solving polynomial equations, *J. ACM*, 8 (1961), 151–162.
- [20] O. L. Rasmussen, Solution of polynomial equations by the method of D. H. Lehmer, *BIT*, 4 (1964), 250–260.

- [21] J. A. Grant and G. D. Hitchins, An always convergent minimization technique for the solution of polynomial equations, *J. Inst. Math. Appl.*, 8 (1971), 122–129.
- [22] J. M. Ortega, *Numerical analysis, a second course*, Academic Press, 1972.
- [23] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, 1970.
- [24] R. Wait, *The numerical solution of algebraic equations*, Wiley, 1979.
- [25] G. R. Garside, P. Jarratt, and C. Mack, A new method for solving polynomial equations, *Comput. J.*, 11 (1968), 87–89.
- [26] B. T. Smith, Error bounds for zeros of a polynomial based upon Gerschgorin's theorems, *J. ACM*, 17 (1970), 661–674.
- [27] A. S. Householder, *The numerical treatment of a single nonlinear equation*, McGraw-Hill, 1970.
- [28] J. H. Wilkinson, *Rounding errors in algebraic processes*, HMSO, 1963.
- [29] J. H. Wilkinson, *Algebraic eigenvalue problem*, Oxford Univ. Press, 1965.
- [30] P. Rabinowitz (ed.), *Numerical methods for nonlinear algebraic equations*, Gordon & Breach, 1970.

302 (XV.4) Numerical Solution of Linear Equations

A. Condition of Linear Systems

The solution of the system of linear algebraic equations

$$\sum_{j=1}^n a_{ij}x_j = b_i; \quad i = 1, \dots, n, \quad a_{ij}, b_i \text{ real}, \quad (1)$$

which may be written in the matrix form

$$Ax = b; \quad A = (a_{ij}), \quad x = (x_i), \quad b = (b_i), \quad (1')$$

is expressed as quotients of determinants by Cramer's rule. In practice, however, this form of the solution is of little value for numerical computation, because the direct evaluation of the determinants involves $(n+1)!(n-1)$ multiplications which, even for a moderate-sized system, amount to a prohibitive number of arithmetic operations to be executed, even with modern high-speed computers; besides, it requires high-precision calculation. There are a variety of practical methods of solving efficiently the system (1) with finite-precision calculation. These numerical methods are generally divided into two classes: direct methods and iterative methods.

Whatever method is used, inherent difficulties are encountered when the solution of the system is unstable. The instability is usually measured by the **condition number** of the coefficient matrix, which is defined by

$$\text{cond } A = \sigma_1 / \sigma_n \quad (\|A\| \|A^{-1}\| \geq 1), \quad (2)$$

where the σ_i ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$) are the non-negative square roots of the eigenvalues of $A^t A$ (called the **singular values** of A), $\|A\| = \max \{\|Ax\|/\|x\| : x \neq 0\}$, and $\|x\| = \sqrt{x^t x}$. This definition is valid also for an $m \times n$ matrix A ($m \geq n$). The condition number $\text{cond } A$ satisfies

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\text{cond } A}{1 - \frac{\|\delta A\|}{\|A\|} \text{cond } A} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right),$$

where $Ax = b$, $(A + \delta A)(x + \delta x) = b + \delta b$, and $\|\delta A\| \|A^{-1}\| < 1$. As the condition number increases, solution processes become more susceptible to errors. If $\text{cond } A$ is large, the system is called **ill-conditioned**.

B. Direct Methods

A direct method is one that yields an exact solution in a finite number of arithmetic operations if they are performed without roundoff error. Among the existing direct methods, the one known as **Gaussian elimination** with pivoting, which is based on systematic elimination of unknowns of the equation (1), is found to be the best with respect to time or accuracy. Its dependability was re-established by means of **backward error analysis** (\rightarrow 138 Error Analysis; [1–3]). The method generates successive vectors $b^{(k)} = (b_i^k)$ and matrices $A^{(k)} = (a_{ij}^k)$, typically of the form

$$A^{(2)} = \begin{pmatrix} * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix}$$

when $n = 5$ and $k = 2$. At the k th stage, a pivotal element $a_{ij}^{k-1} \neq 0$ ($i, j \geq k$) is chosen in one way or another. Then the i th and k th rows and the j th and k th columns of $A^{(k-1)}$ are interchanged so that a_{ij}^{k-1} becomes a_{kk}^k . The i th and k th elements of $b^{(k-1)}$ are also interchanged. The **pivot** a_{kk}^k is then used to eliminate all the nonzero entries in its column below the diagonal as follows:

$$a_{ij}^k = a_{ij}^{k-1}, \quad i = 1, \dots, k, \quad j = 1, \dots, n;$$

$$b_i^k = b_i^{k-1}, \quad i = 1, \dots, k;$$

$$a_{ij}^k = a_{ij}^{k-1} - (a_{ik}^{k-1}/a_{kk}^k) a_{kj}^k,$$

$$i = k + 1, \dots, n, \quad j = k, \dots, n;$$

$$b_i^k = b_i^{k-1} - (a_{ik}^{k-1}/a_{kk}^k) b_k^k, \quad i = k + 1, \dots, n.$$

Starting with $A^{(0)} = A$ and $b^{(0)} = b$, a system of linear equations with an upper-triangular coefficient matrix,

$$A^{(n-1)} x^* = b^{(n-1)},$$

is produced at the stage $k = n - 1$, where x^* is a permutation of x , caused by interchanging columns. This part of the process is called **forward elimination**. If $a_{ij}^{k-1} = 0$ for all $i, j \geq k$ at some stage, then A is a singular matrix of rank $k - 1$ and the system (1) admits infinitely many solutions if $b_i^{k-1} = 0$ for $i = k, \dots, n$, and no solution otherwise. If this is not the case, A is a nonsingular matrix with $\det A = a_{11}^{n-1} a_{22}^{n-1} \dots a_{nn}^{n-1}$, and the solution of (1) is given by

$$x_i^* = \left(b_i^{n-1} - \sum_{k=i+1}^n a_{ik}^{n-1} x_k^* \right) / a_{ii}^{n-1}$$

for $i = n, n - 1, \dots, 1$, which is called **back substitution**. Taking the pivot to be an element of the largest absolute value among column elements a_{ik}^{k-1} ($i \geq k$) at each stage is called the **partial pivoting** strategy. In **complete pivoting**, the pivot is taken to be an element of the largest absolute value among a_{ij}^{k-1} ($i, j \geq k$).

These pivotings are introduced to prevent loss of accuracy due to rapid growth of elements of successive $A^{(k)}$. Although a smaller bound for the growth factor is obtained for complete pivoting, in practice partial pivoting appears to be entirely adequate.

If rows or columns are not interchanged in the process, forward elimination effectively produces a factorization of A into the product of a lower triangular matrix L and an upper triangular matrix U , i.e.,

$$A = LU, \quad (3)$$

where L has unit diagonal elements and $U = A^{(n-1)}$. This factorization is computed directly by the **Doolittle method** without calculating the intermediate $A^{(k)}$. The **Crout method** also produces a similar factorization (3) in which U has unit diagonal elements. The **Cholesky method** determines a similar factorization (3) of a positive definite matrix A , in which $U = L^t$. Once the **triangular factorizations** (3) are formed, the solution of the system (1) is determined by solving the two triangular systems $Ly = b$ and $Ux = y$ successively.

The number of multiplicative operations required for these factorizations are about $n^3/3$ for forward elimination, Doolittle's method, and Crout's method, and about $n^3/6$ for Cholesky's method. The solution of each triangular system requires about $n^2/2$ multiplicative operations. Special properties of A , such as a banded structure, can be exploited to reduce the number of operations and memory requirements considerably [4]. **Gauss-Jordan elimination** is similar to Gaussian elimina-

tion, but the elements above the diagonal are also eliminated to dispense with back substitution. However, it requires about $n^3/2$ multiplications. Modern computers can easily handle problems of size $n = 100$ by these direct methods.

C. Iterative Methods

An iterative method is a dynamical process that generates a sequence of approximations $\{x^k\}$ converging to the exact solution. At each step of the iteration, an improved approximation is obtained from the previous ones. The accuracy of the solution depends on the number of iterations performed. Most iterative methods retain the coefficient matrix in its original form throughout the process and hence have the advantage of requiring minimal memory. They are suitable for solving large **sparse** systems arising in finite-dimensional approximations to \dagger partial differential equations, where A is sparse if most of its elements are zero.

Linear stationary iterative processes are the most frequently used iterative methods. A method in this class is written in the form

$$x^k = x^{k-1} + R(b - Ax^{k-1}), \quad k = 1, 2, \dots, \quad (4)$$

where R is chosen to approximate the inverse of A . Usually, R and $b - Ax^{k-1}$ are not expressed explicitly in the actual algorithms. If the \dagger spectral radius of the **iteration matrix** $I - RA$ is less than one, the method converges to the solution for an arbitrary initial approximation x^0 [5]. The key matrix R can be chosen quite freely as long as this condition is satisfied. In the **Richardson method**, $R = \alpha A^{-1}$, $0 < \alpha < 2/\|A\|^2$. $R = (A + E)^{-1}$ is the usual choice, in which E is a perturbing matrix that makes $A + E$ easily invertible. In the **Gauss-Seidel method** and the **Jacobi method**, R is chosen to be the inverse of the lower triangular and the diagonal submatrices of A , respectively, i.e., $R = (L + D)^{-1}$ and $R = D^{-1}$, where $L = (a_{ij})$ ($i > j$) and $D = (a_{ii})$.

Direct methods can be combined with the iterative method (4) to obtain an approximate inverse R . In the course of factorization by a direct method, an artificial perturbation E is introduced to produce an **incomplete factorization**,

$$\tilde{L}\tilde{U} = A + E, \quad (5)$$

so that \tilde{L} and \tilde{U} are of low \dagger complexity of computation. A combined method is then constructed by putting $R = \tilde{U}^{-1}\tilde{L}^{-1}$. Even if E is not added intentionally, each of the direct methods actually produces an incomplete factorization (5), in which E is a matrix of

small elements accounting for the effect of roundoff errors [1–4]. In this case, the combined method is called the **iterative improvement** of the direct method. If applied to a first solution $x^0 = \tilde{U}^{-1}\tilde{L}^{-1}b$, it produces more accurate solution in a few steps with only a modest increase in computation time, when the system is not too ill-conditioned [4]. It is essential, however, that the residual $b - Ax^0$ be computed with higher precision.

Various convergence criteria have been established for a variety of methods [5, 6]. For example, the Gauss-Seidel method is convergent for a symmetric positive definite matrix A . The smaller the spectral radius, the faster the convergence of the method. In general, a larger amount of computation is required in each iteration to get faster convergence. If the spectral radius $\rho(I - RA)$ is close to one, the convergence is slow, and an acceleration of the process is needed. **SOR** (successive over-relaxation) is an accelerated version of the Gauss-Seidel method, in which $R_\omega = (L + \omega^{-1}D)^{-1}$ and the optimum **acceleration parameter** ω ($0 < \omega < 2$) is chosen to minimize $\rho(I - R_\omega A)$ (\rightarrow 304 Numerical Solution of Partial Differential Equations; [6]). There is an adaptive acceleration method [7], which, if applied to a scalar sequence, reduces to Aitken's δ^2 -method.

D. The Conjugate Gradient Method

The **conjugate gradient (CG) method** is a nonlinear stationary iterative method for solving a system with a symmetric positive definite coefficient matrix. The method generates $\{x^k\}$, $\{r^k\}$, and $\{p^k\}$ by means of the formulas

$$\begin{aligned} \alpha_k &= (r_k, r_k)/(p^k, Ap^k), \\ x^{k+1} &= x^k + \alpha_k p^k, \\ r^{k+1} &= b - Ax^{k+1} = r^k - \alpha_k Ap^k, \\ \beta_k &= (r^{k+1}, r^{k+1})/(r^k, r^k), \\ p^{k+1} &= r^{k+1} + \beta_k p^k, \end{aligned} \quad (6)$$

where x^0 is arbitrary and $p^0 = r^0 = b - Ax^0$, $(r, s) = s^T r$.

The CG method shares a feature with the direct method. In theory, $\{x^k\}$ converges to the solution in less than n steps, and the p^i are mutually **conjugate**, i.e., $(p^i)^T Ap^j = 0$ ($i \neq j$). When Hestenes and Stiefel [9] proposed the method in 1952, they created a great sensation because of the method's theoretical elegance. The CG method turned out, however, to be highly sensitive to roundoff errors. In practice, nice theoretical properties, such as finite termination, do not hold in the presence of error. Recently, the CG method has regained its

popularity as an iterative method for solving large sparse systems. The iteration (6) is usually restarted periodically to accelerate the convergence. It may converge in a small number of steps, as it takes advantage of the distribution of eigenvalues of A adaptively. The CG method is most effective when it is used to accelerate linear stationary iterative methods or when it is applied to the **preconditioned** system

$$\tilde{L}^{-1}A(\tilde{L}^{-1})^T y = L^{-1}b, \quad x = (L^{-1})^T y, \quad (7)$$

where \tilde{L} is computed by the incomplete Cholesky factorization $\tilde{L}\tilde{L}^T = A + E$. The matrix $\tilde{L}^{-1}A(\tilde{L}^{-1})^T$ is not to be formed explicitly in the CG method.

E. Linear Least Squares Problem

Let $Ax = b$ be an overdetermined system of linear equations, where A is an $m \times n$ matrix ($m \geq n$). The **linear least squares problem** is to find the x that minimizes the Euclidean norm $\|b - Ax\|$. We assume here that A is of rank n . The most straightforward method for solving the problem is to apply the Cholesky method to the **normal equation**

$$(A^T A)x = A^T b. \quad (8)$$

This may result in ill-conditioning of the system, since $\text{cond } A^T A = (\text{cond } A)^2$. Ill-conditioning can sometimes be avoided by forming the factorization $A = LU$ by Gaussian elimination, where L is an $m \times n$ lower trapezoidal matrix and U is an upper triangular matrix. The least squares solution is then obtained by solving successively the systems $L^T L y = L^T b$ and $Ux = y$.

Another approach to avoiding ill-conditioning is based on orthogonalization. The modified Gram-Schmidt orthogonalization method produces in an effective manner the factorization

$$A = QU, \quad (9)$$

where Q is an $m \times n$ matrix whose columns are a set of orthonormal vectors and U is an upper triangular matrix [10, 11]. The least squares solution is obtained by solving $Ux = Q^T b$. A sequence of **Householder transformations**, $P_k = I - 2w_k w_k^T / \|w_k\|^2$, can produce the factorization

$$P_n P_{n-1} \dots P_2 P_1 A = \begin{pmatrix} U \\ 0 \end{pmatrix}, \quad (10)$$

where U is an $n \times n$ upper triangular matrix [12]. The least squares solution is obtained by solving $Ux = \bar{b}$, where \bar{b} is formed by the first n elements of the vector $P_n P_{n-1} \dots P_2 P_1 b$. A

similar factorization can also be produced by using a sequence of **Givens transformations**, which have the advantage of exploiting the sparseness of A . The matrices U in (9) and (10) are unique and coincide with the transpose L^T of the lower triangular matrix L of the Cholesky factorization, $A^T A = LL^T$, of $A^T A$. The multiplicative operations required for the factorizations (9) and (10) are about mn^2 and $mn^2 - n^3/3$, respectively.

When the rank of the matrix is unknown, we can determine an "effective rank" p of A , based on the **singular value decomposition** (SVD):

$$A = UDV^T,$$

where U and V are orthogonal matrices of dimensions m and n , respectively, and D is an $m \times n$ diagonal matrix whose diagonal elements d_{ii} are the singular values σ_i of A [13]. If singular values smaller than σ_p can be ignored, the approximate least squares solution is given by $\hat{x}^p = V \tilde{D}_p^+ U^T b$, where the i th element \tilde{d}_{ii}^p of the diagonal matrix \tilde{D}_p^+ is σ_i^{-1} if $i \leq p$ and 0 otherwise.

These methods certainly solve the system (1) but generally require more computational work than those based on the triangular factorization. The least squares solution is computed also by applying the CG method (6) to the normal equation (8), in which $A^T A$ is not to be formed explicitly [9]. Iterative methods for solving linear systems with singular and rectangular coefficient matrices are characterized in terms of the range of RA and the null space of AR [15].

Besides the methods discussed here, numerous other methods have been proposed for solving the system (1) [16–18]. Recently, various sparse techniques have been developed for direct methods to control the growth of non-zero entries (**fill-in**) in the process of matrix factorization [19]. In general, direct methods have an advantage over iterative methods; however, their relative computational efficiency varies according to the scale, sparsity, and type of the coefficient matrix and to the available computational devices.

References

- [1] J. H. Wilkinson, Error analysis of direct methods of matrix inversion, *J. ACM*, 8 (1961), 281–330.
- [2] J. H. Wilkinson, *Rounding errors in algebraic processes*, Prentice-Hall, 1963.
- [3] J. H. Wilkinson, *The algebraic eigenvalue problem*, Oxford Univ. Press, 1965.
- [4] G. E. Forsythe and C. B. Moler, *Computer solution of linear algebraic systems*, Prentice-Hall, 1967.

- [5] R. S. Varga, Matrix iterative analysis, Prentice-Hall, 1962.
- [6] D. M. Young, Iterative solution of large linear systems, Academic Press, 1971.
- [7] K. Tanabe, An adaptive acceleration of general linear iterative process for solving a system of linear equations, Proc. Fifth Hawaii Int. Conf. System Sci., (1972), 116–118.
- [8] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat. Bur. Standards, sec. B, 45 (1950), 255–282.
- [9] M. R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Standards, sec. B, 49 (1952), 409–436.
- [10] F. L. Bauer, Elimination with weighted row combinations for solving linear equations and least squares problems, Numer. Math., 7 (1965), 338–352.
- [11] A. Björck, Solving linear least squares problems by Gram-Schmidt orthogonalization, BIT, 7 (1967), 1–21.
- [12] G. H. Golub, Numerical methods for solving linear least squares problems, Numer. Math., 7 (1965), 206–216.
- [13] G. Golub and W. Kahan, Calculating the singular values and pseudoinverse of a matrix, SIAM J. Numer. Anal., 2 (1965), 205–224.
- [14] C. L. Lawson and R. J. Hanson, Solving least squares problems, Prentice-Hall, 1974.
- [15] K. Tanabe, Characterization of linear stationary iterative processes for solving a singular system of linear equations, Numer. Math., 22 (1974), 349–359.
- [16] D. K. Faddeev and V. N. Faddeeva, Computational methods of linear algebra, Freeman, 1963. (Original in Russian, 1963.)
- [17] G. E. Forsythe, Solving linear algebraic equations can be interesting, Bull. Amer. Math. Soc., 59 (1953), 299–329.
- [18] J. R. Westlake, A handbook of numerical matrix inversion and solution of linear equations, Wiley, 1968.
- [19] J. K. Reid (ed.), Large sparse sets of linear equations, Academic Press, 1971.

303 (XV.8) Numerical Solution of Ordinary Differential Equations

A. General Remarks

Applications of the numerical solution of *ordinary differential equations include *initial value problems, *boundary value problems,

and *eigenvalue problems. Although solution by an *analog computer is among the numerical methods in the wide sense, we usually mean by “numerical solution” a solution obtained by approximating a problem of infinite degrees of freedom and a continuous variable by one of finite degrees of freedom. The approximation of an arbitrary function by a linear combination of a given finite system of functions is an example (\rightarrow Section I). Successive substitution and power series expansion are also used but have no particular advantages over other methods. Among various methods of approximation, the so-called **difference methods** (or **discrete variable methods**) are the most flexible and have the largest field of application. A difference method reduces a given problem to an approximate problem in which we deal only with a systematically chosen set of discrete values of the variable and with values of the unknown functions at the chosen points. A difference method is usually carried out by a simple iterative computation, so that it is suitable for *digital computers. We confine ourselves almost exclusively to difference methods [1–7].

B. Initial Value Problems

To solve numerically general initial value problems, it suffices to consider the problem of determining numerically the values on an interval $[a, b]$ of m functions $y^i(x)$ ($i = 1, \dots, m$) that satisfy a system of ordinary differential equations of the form $y^{i\prime\prime}(x) = f^i(x, y^1(x), \dots, y^m(x))$ and the initial conditions $y^i(a) = \eta^i$ ($i = 1, \dots, m$), where the f^i are given functions with appropriate smoothness, and a, b , and η^i are given constants. We define the mesh points $x_n = a + nh$ ($n = 0, 1, \dots$), call h the **step size**, and determine numerically the values y_n^i approximating $y^i(x_n)$.

Assuming that the y_n^i were computed with infinite accuracy, i.e., without *roundoff error, we call $e_n^i \equiv y_n^i - y^i(x_n)$ the global **truncation error** (or global **discretization error**), while we call $r_n^i \equiv \tilde{y}_n^i - y_n^i$ the global **roundoff error**, where the \tilde{y}_n^i are the values we actually have by means of finite-accuracy computation. On the other hand, if the assumption holds locally that the solutions at the previous steps are exact, we call the e_n^i and r_n^i the local truncation error (or local discretization error) and the local roundoff error, respectively. To avoid confusion, we denote the local truncation error by t_n^i . The numerical solution must have the property that, for every $x \in [a, b]$, $\lim_{h \rightarrow 0, x_n = x} e_n^i = 0$, which is called the condition of **convergence**. In particular, if $t_n^i = O(h^{p+1})$ ($x_n = x, h \rightarrow 0$, where $O(h^{p+1})$ is Landau's symbol), then p

is called the **order** of the solution process.

We shall show the typical methods of numerical solution, using the abbreviation f_n^i for $f^i(x_n, y_n^i)$.

C. Overall Approximation

If we rewrite the differential equations on the interval $[x_n, x_{n+P}]$ as $y^i(x_{n+q}) - y^i(x_n) = \int_{x_n}^{x_{n+q}} f^i(x, y^j(x)) dx$, $q = 1, \dots, P$, and substitute suitable numerical integration formulas for the integrals on the right-hand side, then we have the **overall approximation formulas**:

$$y_{n+q}^i - y_n^i = h \sum_{r=0}^P c_{qr} f_{n+r}^i, \quad q = 1, \dots, P,$$

from which we can obtain $y_{n+1}^i, \dots, y_{n+P}^i$ when the y_n^i are given. Since f_{n+r}^i contains y_{n+r}^i , these equations are nonlinear in y_{n+q}^i , so that we have to use, for example, the iterative procedure of starting from suitable 0th approximations and adopting as the l th approximations of the y_{n+q}^i values computed from the overall approximation formulas by substituting the $(l-1)$ th approximations for the y_{n+r}^i in the f_{n+r}^i on the right-hand sides. These formulas are often used to determine a set of starting values for a multistep method. The truncation error of the formulas depends on that of the numerical integration formula substituted for the integral. A few examples: $P = 1$, $(c_{10}, c_{11}) = (1/2, 1/2)$ (error $h^3 y^{(3)}(x_n)/12 + O(h^4)$); $P = 2$, $(c_{10}, c_{11}, c_{12}) = (5/12, 8/12, -1/12)$ (error $-h^4 y^{(4)}(x_n) + O(h^5)$), $(c_{20}, c_{21}, c_{22}) = (1/3, 4/3, 1/3)$ (error $h^5 y^{(5)}(x_n) + O(h^6)$); $P = 3$, $(c_{10}, c_{11}, c_{12}, c_{13}) = (9/24, 19/24, -5/24, 1/24)$, $(c_{20}, c_{21}, c_{22}, c_{23}) = (1/3, 4/3, 1/3, 0)$, $(c_{30}, c_{31}, c_{32}, c_{33}) = (3/8, 9/8, 9/8, 3/8)$.

D. Runge-Kutta Methods

By a **general Runge-Kutta method** we mean a method of determining $y_0^i = \eta^i, y_1^i, y_2^i, \dots$ successively by means of a formula $y_{n+1}^i - y_n^i = h\Phi^i(x_n, y_n^j, h)$, where the functions Φ^i are defined in terms of parameters α_r and β_{rs} as $\Phi^i(x, y^j, h) = \sum_{r=0}^P \alpha_r k_r^i$, $k_0^i = f^i(x, y^j)$, $k_r^i = f^i(x + \beta_{r0}h, y^j + h(\sum_{s=1}^P \beta_{rs} k_s^i + h \sum_{s=1}^P \beta_{rs} k_s^j))$ ($r = 1, \dots, P$). A Runge-Kutta method is called **explicit** if $\beta_{rs} = 0$ for $r \leq s$, and **implicit** otherwise. In the latter case, if $\beta_{rs} = 0$ for $r < s$, the method is called **semi-explicit** (or **semi-implicit**). Unless otherwise stated, the Runge-Kutta methods considered below are assumed to be explicit. In order to proceed one step from y_n^i to y_{n+1}^i we have to compute each function f^i ($P+1$) times. Hence this is called the **$(P+1)$ -stage method**. If we denote by $z^i(t)$ the solutions

of $z^{i'}(t) = f^i(t, z^j(t))$ with the initial conditions $z^i(x) = y^i$ and put $h\Delta^i(x, y^j; h) = z^i(x+h) - z^i(x)$, then we have $\Phi^i(x, y^j; h) - \Delta^i(x, y^j; h) = h^p \varphi^i(x, y^j) + O(h^{p+1})$, where the $\varphi^i(x, y^j)$ are expressible in terms of the $f^i(x, y^j)$ and their (partial) derivatives. Various Runge-Kutta methods have been devised by searching for values of the α_r and β_{rs} that make p as large as possible for a given p (p is called the order of the method). When searching for these values of α_r and β_{rs} , we usually impose at least one of the following conditions: (1) α_r and β_{rs} are simple; (2) the truncation error is small; (3) the region of absolute stability (\rightarrow Section G) is large; (4) α_r and β_{rs} give smaller roundoff errors; (5) only a small computer memory is required. The ordered pairs (P, p) , where p is the highest order that can be attained by the $(P+1)$ -stage method, are $(P \leq 3, P+1)$, $(4, 4)$, $(5, 5)$, $(6, 6)$, $(7, 6)$, $(8, 7)$, $(9-10, 8)$, $(p+2 \leq (P+1) \leq (1/2)(p^2 - 2p + 4), p \geq 9)$. The following formulas are frequently used since they are accurate and have simple parameters: for $P = 1$ and $p = 2$, the formulas with $\alpha_0 = 0, \alpha_1 = 1, \beta_{10} = 1/2$ (modified Euler method), with $\alpha_0 = \alpha_1 = 1/2, \beta_{10} = 1$ (improved Euler method) and with $\alpha_0 = 1/4, \alpha_1 = 3/4, \beta_{10} = 2/3$ (Ralston's second-order method); for $P = 2$ and $p = 3$, the formulas with $\alpha_0 = 2/9, \alpha_1 = 1/3, \alpha_2 = 4/9, \beta_{10} = 1/2, \beta_{20} = 3/4, \beta_{21} = 3/4$ (Ralston's third-order method), and with $\alpha_0 = 1/4, \alpha_1 = 0, \alpha_2 = 3/4, \beta_{10} = 1/3, \beta_{20} = 2/3, \beta_{21} = 2/3$ (Heun's third-order method). For $P = 3, p = 4$, the formula with $\alpha_0 = \alpha_3 = 1/6, \alpha_1 = \alpha_2 = 1/3, \beta_{10} = \beta_{20} = \beta_{21} = 1/2, \beta_{30} = \beta_{32} = 1, \beta_{31} = 0$ is well known and is frequently referred to as "the fourth-order Runge-Kutta method" or "the Runge-Kutta method." It has various desirable features [8]. Gill's modification of the classical Runge-Kutta method (sometimes called the **Runge-Kutta-Gill method**) has some advantage in regard to roundoff errors and to the necessary computer memory size, while Ralston's second-, third-, and fourth-order methods have minimum error bounds in Lotkin's sense for the local truncation error [7]. There exist "substantially fifth-order" methods for $P = 4$ [9]. Today we frequently prefer to use higher-order methods, especially fifth-order methods, instead of lower-order methods, because the former require less computation time for solutions of given accuracy [10].

For the estimation of the local truncation error, the following two practical techniques are well known: (i) **one-step-two-half-steps error estimate**. If we let y_{n+1}^{12} and y_{n+1}^i denote the approximations that are computed with an m th-order method by taking two half-steps and one full step ($= h_n$), respectively, then an estimate of the local truncation error per half-step associated with y_{n+1}^{12} is given by the

formula

$$t_{n+1}^i \approx \gamma_m^i (h_n/2)^{m+1} \approx (y_{n+1}^{i2} - y_{n+1}^i)/2(1-2^m),$$

where we assume that γ_m^i does not vary much over the interval $[x_n, x_{n+1}]$. (ii) **The formula of embedding form.** The following formula of embedding form is more effective than the traditional method above. By use of the p th-order method $y_{n+1}^i = y_n^i + \sum_{r=0}^p \alpha_r k_r^i$ and the $(p+1)$ th-order method $y_{n+1}^{i*} = y_n^i + \sum_{r=0}^p \alpha_r^* k_r^i$, we obtain an estimate of the local truncation error in y_{n+1}^i as $t_{n+1}^i \approx y_{n+1}^i - y_{n+1}^{i*}$ [11–13]. At present, formulas of this type are known for $p = 2-8$, where the method for $p=5$ is said to be the most efficient. Recently, a similar method for the estimation of the global truncation error has been investigated [14]. In addition, multistep generalizations of explicit Runge-Kutta methods, which are known as **pseudo-Runge-Kutta methods**, have also been investigated [15].

E. Multistep Methods

A **linear multistep method** approximates given differential equations by difference equations $\rho(E)y_n^i = h\sigma(E)f_n^i$, where E is the operator of increasing n by 1, $\rho(\zeta) = \alpha_k \zeta^k + \alpha_{k-1} \zeta^{k-1} + \dots + \alpha_1 \zeta + \alpha_0$, $\sigma(\zeta) = \beta_k \zeta^k + \beta_{k-1} \zeta^{k-1} + \dots + \beta_1 \zeta + \beta_0$, $\alpha_k \neq 0$, and $|\alpha_0| + |\beta_0| \neq 0$. (If ρ and σ are of degree k in ζ , we speak of a **linear k -step method**.) By means of these difference equations we can determine y_{n+k}^i from $y_{n+k-1}^i, y_{n+k-2}^i, \dots, y_n^i$. Since the y_{n+k}^i are determined explicitly if $\beta_k = 0$ and implicitly if $\beta_k \neq 0$, the difference equations are called **explicit** and **implicit**, respectively. In the implicit case, if $|h| < |(\alpha_k/\beta_k)L|$, where L is a Lipschitz constant for $f^i(x, y^j)$, then the y_{n+k}^i can be calculated by successive substitutions. In order to obtain the y_n^i successively by means of a k -step method with $k \geq 0$, it is necessary, to begin with, to give the $k-1$ sets of m values y_1^i, \dots, y_{k-1}^i (called the **starting values**) in addition to the initial values $y_0^i = \eta^i$. To determine the starting values, Runge-Kutta methods, overall approximation formulas, etc., are ordinarily used. The number of times we have to compute the f^i to proceed one step from y_n^i to y_{n+1}^i is only 1 for an explicit case, while for an implicit case it is equal to the number of iterations required to secure the convergence of the successive substitutions (ordinarily, the step size h as well as the 0th approximation for y_{n+k}^i are chosen so that the convergence is attained after a few iterations). We have $\lim_{h \rightarrow 0, x_n = x} e_n^i = 0$ ($x \in [a, b]$) for any f^i and η^i and any starting values such that $\lim_{h \rightarrow 0} y_\mu^i = \eta^i$ ($\mu = 0, 1, \dots, k-1$), if and only if the polynomials ρ and σ satisfy the following two conditions: (i) **consistency**: $\rho(1) = 0$, $\rho'(1) = \sigma(1)$; (ii) **zero stability**: Every root

of $\rho(\zeta) = 0$ lies on or inside the unit circle, and every root lying on the unit circle is simple.

We call $t_n^i = (\rho(E)y^i(x_n) - h\sigma(E)y^{i*}(x_n))/\sigma(1)$ the **local truncation error**. For a consistent and stable method determined by (ρ, σ) there exist a real constant C_{p+1} ($\neq 0$) and an integer p (≥ 1) such that $t_n^i = h^{p+1} C_{p+1} y^{i(p+1)}(x_n)/\sigma(1) + O(h^{p+2})$; p is called the **order** of the method and $C \equiv C_{p+1}/\sigma(1)$ the **error constant**. For a given ρ satisfying (ii), σ is chosen to make the order as small as possible and then to make the error constant as small as possible. However, p cannot exceed $k+1$ if k is odd or $k+2$ if k is even; moreover, p can be equal to $k+2$ only if all the roots of $\rho(\zeta) = 0$ lie on the unit circle.

The following are examples of multistep methods.

(1) Explicit Methods.

(a) Adams-Bashforth methods.

$$k=1, \quad \rho(\zeta) = \zeta - 1, \quad \sigma(\zeta) = 1, \quad p=1,$$

$$C = 1/2 \quad (\text{Euler method});$$

$$k=2, \quad \rho(\zeta) = \zeta(\zeta - 1), \quad \sigma(\zeta) = (3\zeta - 1)/2,$$

$$p=2, \quad C = 5/12;$$

$$k=3, \quad \rho(\zeta) = \zeta^2(\zeta - 1),$$

$$\sigma(\zeta) = (23\zeta^2 - 16\zeta + 5)/12, \quad p=3,$$

$$C = 3/8;$$

etc.

(b) Midpoint rule.

$$k=2, \quad \rho(\zeta) = \zeta^2 - 1, \quad \sigma(\zeta) = 2\zeta,$$

$$p=2, \quad C = 1/6.$$

(c) Milne's predictor.

$$k=4, \quad \rho(\zeta) = \zeta^4 - 1,$$

$$\sigma(\zeta) = (8\zeta^3 - 4\zeta^2 + 8\zeta)/3,$$

$$p=4, \quad C = 7/90.$$

(2) Implicit Methods.

(a) Adams-Moulton methods.

$$k=1, \quad \rho(\zeta) = \zeta - 1, \quad \sigma(\zeta) = (\zeta + 1)/2,$$

$$p=2, \quad C = -1/12 \quad (\text{trapezoidal rule});$$

$$k=2, \quad \rho(\zeta) = \zeta(\zeta - 1),$$

$$\sigma(\zeta) = (5\zeta^2 + 8\zeta - 1)/12,$$

$$p=3, \quad C = -1/24;$$

etc.

(b) Milne's corrector (or the Milne-Simpson formula).

$$k=2, \quad \rho(\zeta) = \zeta^2 - 1, \quad \sigma(\zeta) = (\zeta^2 + 4\zeta + 1)/3,$$

$$p=4, \quad C = -1/90.$$

When an implicit formula (ρ, σ) is used to obtain y_{n+k}^i , the 0th approximation y_{n+k}^{i*} for y_{n+k}^i is usually determined by an explicit formula (ρ^*, σ^*) of the same order as (ρ, σ) (where (ρ^*, σ^*) itself may not necessarily be stable). This kind of combination of implicit and explicit methods is called a **predictor-corrector method** (or **PC method**), where the formula (ρ^*, σ^*) is called the **predictor** and (ρ, σ) the **corrector**. Typical combinations are the midpoint rule and the trapezoidal rule, Milne's predictor and Milne's corrector (this combination is called **Milne's method**), and an Adams-Bashforth method and an Adams-Moulton method of the same order. A predictor-corrector method has the advantage that the local truncation error (of the corrector) can be estimated in the course of calculation without extra computation. In fact, if the orders of the predictor (ρ^*, σ^*) and the corrector (ρ, σ) are equal to p and their error constants to C^* and C , respectively, then the local truncation error t_n^i can be estimated by $t_n^i = KD_n^i + O(h^{p+2})$ ($n = 0, 1, \dots$), where $D_n^i \equiv y_{n+k}^{i*} - y_{n+k}^i$ is the difference between the value y_{n+k}^{i*} at x_{n+k} obtained by the predictor and the y_{n+k}^i obtained by the corrector, and $K = \alpha_k^* C / [(C - C^*)\sigma^*(1)]$ [16].

The term **multistep methods** originates from the fact that these use the values of the dependent variables at more than two different mesh points in order to proceed one step. These are also called **multivalue methods** since they use more than one value of the dependent variable. The multivalue method is, however, a more general concept than the multistep method.

Linear multistep methods are not only examples of the PC method; they are also examples of the **variable-step variable-order algorithms** (VSVO algorithms), where the order of the formula as well as the step size are automatically chosen according to the behavior of the solution. In practical VSVO algorithms, the Adams-Bashforth-Moulton family of PC pairs of order 1 to 13 are usually used, and for solving stiff systems (\rightarrow Section G) those correctors with large regions of absolute stability are used. In these algorithms we use the multivalue method, which saves the information at different steps in a form convenient for the change of order and of step size. The multivalue method also facilitates error estimation and stability, and results in an efficient use of memory and reduction of the computational cost. For example, in Gear's algorithm the information required for computing y_{n+1} is saved in the following form: $\mathbf{y}_n = (y_n, hy_n', \dots, (h^k/k!)y_n^{(k)})$, where $y'(x)$ is the polynomial $P_{k,n}(x)$ interpolating $f_n, f_{n-1}, \dots, f_{n-k+1}$, where $y_n^{(i)} = d^{i-1}P_{k,n}(x)/dx^{i-1}|_{x=x_n}$ hold. All the local truncation error estimators are based essentially

on Milne's device. In the VSVO algorithms heuristics play an important role [2, 17, 18].

F. Extrapolation Methods

Let us assume that the numerical solution with the step size h at some fixed point x has the form

$$y(x, h) = y(x) + \sum_{i=1}^m \tau_i h^r + O(h^{r(m+1)}),$$

where r is a positive integer. Then we can approximate $y(x, h)$ by a function $R_m(x, h)$ with $(m+1)$ unknowns determined by the requirement that $R_m(x, h_j) = y(x, h_j)$, $j = 0, 1, 2, \dots, m$, $h_i > h_j$ ($i < j$) in order to approximate $y(x)$ by $R_m(x, 0)$. If $R_m(x, h)$ is a polynomial of degree mr in h , we call the foregoing method a **polynomial extrapolation method**, where $R_m(x, h) = y(x) + o(h^{r(m+1)})$ as h tends to 0. Defining $R_0^i = y(x, h_j)$ and $R_m(x, 0) = R_m^i$ for $j = i, i+1, \dots, i+m$, we can calculate the R_m^i by the recursion relation $R_m^i = R_{m-1}^{i+1} + (R_{m-1}^{i+1} - R_{m-1}^i) / ((h_i/h_{i+m})^r - 1)$, $m \geq 1$. It is well known that, for the following method (Gragg's method), there exists an asymptotic expansion with $r = 2$ [19]: $x_n = nh$, $\eta(0, h) = y(0)$, $\eta(x_1, h) = y_0 + hf(y_0, 0)$, $\eta(x_{n+1}, h) = \eta(x_n, h) + 2hf(\eta(x_n, h), x_n)$, $n = 1, 2, \dots, N-1$, where $x_N = x$, $y(x, h) = 1/2[\eta(x_{N-1}, h) + \eta(x_N, h) + hf(\eta(x_N, h), x_N)]$. If $R_m(x, h)$ is a rational function

$$R_m(x, h_j) = (p_0 + p_1 h + p_2 h^2 + \dots + p_n h^n) / (q_0 + q_1 h + q_2 h^2 + \dots + q_v h^v),$$

where $\mu = [m/2]$, $v = m - \mu = [(m+1)/2]$, $j = i, i+1, \dots, i+m$, we call this method a **rational extrapolation method**. For $r = 2$, the following formulas were derived by Bulirsch and Stoer [20]:

$$R_m^i = R_{m-1}^{i+1} + (R_{m-1}^{i+1} - R_{m-1}^i) / ((h_i/h_{i+m})^2 \times [1 - (R_{m-1}^{i+1} - R_{m-1}^i) / (R_{m-1}^{i+1} - R_{m-1}^2)] - 1),$$

$$m \geq 1, R_{-1}^i = 0, R_0^i = y(t, h_i).$$

G. Stability

Consider the application of the general k -step method at the n th point (which is consistent and 0-stable)

$$\sum_{j=0}^k \alpha_j y_{n+j} = h \sum_{j=0}^k \beta_j f_{n+j}$$

to $y' = f(x, y)$, $y(x_0) = y_0$. Let \tilde{y}_n be the numerical solution at $x = x_n$. Let $\tilde{e}_n = y(x_n) - \tilde{y}_n$ be the global error, and let φ_n be the total error at the n th application. Then we find

$$\sum_{j=0}^k \alpha_j \tilde{e}_{n+j} = h \sum_{j=0}^k \frac{\partial f}{\partial y}(x_{n+j}, \xi_{n+j}) \beta_j \tilde{e}_{n+j} + \varphi_{n+k},$$

where ξ_{n+j} lies in the open interval whose endpoints are y_{n+j} and $y(x_{n+j})$. If we make two assumptions, $\partial f/\partial y = \lambda$ (const), $\varphi_n = \varphi$ (const), the above equation reduces to $\sum_{j=0}^k (\alpha_j - h\lambda\beta_j)\tilde{e}_{n+j} = \varphi$, whose general solution is given by

$$\tilde{e}_n = \sum_{s=0}^k d_s r_s^n - \varphi/h\lambda \sum_{j=0}^k \beta_j,$$

where the d_s are arbitrary constants and the r_s are the roots, assumed distinct, of the polynomial equation

$$\sum_{j=0}^k (\alpha_j - h\lambda\beta_j)r^j = 0.$$

We call the linear k -step method **absolutely stable** for a given $h\lambda$ if $|r_s| < 1$, $s = 1, 2, \dots, k$. On the other hand, we call the linear k -step method **relatively stable** for a given $h\lambda$ if $|r_s| < |r_1|$ (or $|r_s| < e^{\lambda h}$), $s = 2, 3, \dots, k$, where r_1 is the root corresponding to the theoretical solution. We call the region $S_a = \{h\lambda \mid |r_s| < 1, s = 1, 2, \dots, k\}$ (or $S_r = \{h\lambda \mid |r_s| < e^{\lambda h} \text{ (or } |r_1|), s = 2, 3, \dots, k\}$) in the complex plane the **region of absolute stability** (or the **region of relative stability**) of the linear k -step method. Also we call $S_a \cap R$ (or $S_r \cap R$) the **interval of absolute stability** (or the **interval of relative stability**) of the linear multistep method, where R is the real line. The explicit methods and the PC methods have a finite interval of absolute stability. The implicit methods usually have larger intervals of absolute stability than the corresponding explicit methods. The higher-order PC methods have smaller intervals, while Runge-Kutta methods do not. The PECE methods usually have larger intervals of stability than the corresponding PEC or P(EC)² methods, where P indicates an application of predictor, C an application of corrector, and E an evaluation of f [21, 22].

Let us consider the linear system $y' = Ay + \varphi(x)$, where A is an $m \times m$ constant matrix and $y(x)$, $y'(x)$, $\varphi(x) \in \mathbf{R}^m$. If A possesses m distinct eigenvalues $\lambda_t = \mu_t + i\nu_t$, $t = 1, 2, \dots, m$, the theoretical solution of this system is given by

$$y(x) = \sum_{t=1}^m K_t \exp((\mu_t + i\nu_t)x) C_t + \psi(x),$$

where K_t and C_t , $t = 1, 2, \dots, m$, are, respectively, arbitrary constants and the eigenvectors corresponding to λ_t and $\psi(x)$ is a particular solution. We call the linear system **stiff** if (i) $\mu_t < 0$, $t = 1, 2, \dots, m$, (ii) $s \gg 1$, where $s = \max_{1 \leq t \leq m} |\mu_t| / \min_{1 \leq t \leq m} |\mu_t|$. We call the non-linear system $y' = f(x, y)$, $y(x)$, $y''(x)$, $f(x, y) \in \mathbf{R}^m$ **stiff in an interval I of x** if, for every $x \in I$, the eigenvalues $\lambda_t(x)$ of the Jacobian matrix of f satisfy (i) and (ii). The ratio s is called the **stiffness ratio**. To solve stiff systems effectively a variety of methods with infinite stability

regions have been proposed. Some of them are:

- (1) **A-stability** (Dahlquist): $S_a \supseteq \{h\lambda \mid \operatorname{Re}(h\lambda) < 0\}$
- (2) **Stiff-stability** (Gear): $S_a \supseteq S_1 \cup S_2$, $S_1 = \{h\lambda \mid \operatorname{Re}(h\lambda) < -a < 0\}$, $S_2 = \{h\lambda \mid -a \leq \operatorname{Re}(h\lambda) \leq b, -c \leq \operatorname{Im}(h\lambda) \leq c, b \geq 0, c > 0\}$
- (3) **A(α)-stability** (Widlund): $S_a \supseteq \{h\lambda \mid -\alpha < \pi - \arg(h\lambda) < \alpha, \alpha \in (0, \pi/2)\}$
- (4) **A₀-stability** (Cryer): $S_a \supseteq \{h\lambda \mid \operatorname{Re}(h\lambda) < 0, \operatorname{Im}(h\lambda) = 0\}$.

The order p and the step number k of linear multistep methods are restricted by the following stability requirements:

- (i) **A-stability**: implicit, $p \leq 2$; trapezoidal rule is the most accurate method.
- (ii) **Stiff-stability**: implicit, $p \leq k$; backward differentiation formulas $\sum_{j=0}^k \alpha_j y_{n+j} = hf_{n+k}$ are stiffly stable for $p = k = 1, 2, \dots, 6$, 0-unstable for $k = 6$.
- (iii) **A(α)-stability**: implicit; there exist high-order A(0)-stable linear multistep methods. (A method is said to be **A(0)-stable** if it is A(α)-stable for some sufficiently small $\alpha \in (0, \pi/2)$.)

For the Runge-Kutta (P, p) methods, we can write

$$\tilde{e}_{n+1} = \left\{ 1 + h\lambda + (h\lambda)^2/2! + \dots + (h\lambda)^p/p! + \sum_{q=p+1}^{P+1} \gamma_q (h\lambda)^q \right\} \tilde{e}_n + \varphi_{n+1},$$

where $\partial f/\partial y = \lambda$ (const) and γ_q are functions of the coefficients of the method in use. We call a region $R = \{h\lambda \mid |1 + h\lambda + (h\lambda)^2/2! + \dots + (h\lambda)^p/p! + \sum_{q=p+1}^{P+1} \gamma_q (h\lambda)^q| < 1\}$ a **region of absolute stability of the Runge-Kutta (P, p) method**.

The implicit Runge-Kutta methods have a larger region of absolute stability than the corresponding explicit Runge-Kutta methods. Therefore the implicit Runge-Kutta methods are suitable for stiff equations, and the explicit Runge-Kutta methods are suitable for nonstiff or mildly stiff equations.

Recently, Yamaguti and others have pointed out that the instabilities occurring in the numerical solution of ordinary differential equations are closely connected to the phenomena of **chaos** studied by Li, Yorke, and others [23, 24].

H. Boundary Value Problems

A boundary value problem is generally formulated as the problem of obtaining functions $y^i(x)$ that satisfy the differential equations $y^{(i)}(x) = f^i(x, y^1(x), \dots, y^m(x))$ and boundary conditions $B_k(y^1(x_{n_1}), \dots, y^m(x_{n_1}); y^1(x_{n_2}), \dots, y^m(x_{n_2}); \dots; y^1(x_{n_s}), \dots, y^m(x_{n_s})) = 0$ ($i, k = 1, \dots, m$), where the f^i and B_k are given functions. If the f^i and B_k are linear in the y^i , we can

first calculate the solutions $y_0^i(x)$ of the differential equations under an appropriate initial condition, as well as a set of m independent solutions $y_l^i(x)$ ($l = 1, \dots, m$) of the homogeneous differential equations (say, under the initial conditions $y_l^i(a) = \delta_l^i$ at some point $x = a$), and then substitute the expression for the desired solution $y^i(x) = y_0^i(x) + \sum_{l=1}^m \alpha_l y_l^i(x)$ in the boundary conditions to obtain simultaneous linear equations for the unknowns α_l (\rightarrow 302 Numerical Solution of Linear Equations). No universally powerful method is known for the case where the f^i or the B_k or both are nonlinear. Ordinarily, we resort to the trial-and-error method of solving the differential equations iteratively under different initial conditions until the solution fully satisfies the boundary conditions, or to approximating the differential operators by suitable difference operators to obtain a set of (generally nonlinear) simultaneous equations approximating at the same time the differential equations and the boundary conditions. For related topics \rightarrow 298 Numerical Computation of Eigenvalues; 315 Ordinary Differential Equations (Boundary Value Problems); [25–27].

I. Methods Other than Difference Methods

Besides difference methods there are frequently used methods for finding in a given finite-dimensional function space a function that best satisfies the differential equations as well as the initial or boundary conditions. Denote the equations by $L[y(x)] = 0$ and assume the conditions to be linear. We choose a function $y_0(x)$ that satisfies the conditions and that is considered to approximate the exact solution, and also functions $y_l(x)$ ($l = 1, \dots, q$) that are considered to represent the typical deviations of $y_0(x)$ from the exact solution and each of which satisfies the homogeneous conditions. We then determine the α_l in such a way that $y(x) = y_0(x) + \sum_{l=1}^q \alpha_l y_l(x)$ best satisfies the differential equations. Corresponding to different interpretations of the words “best satisfy” there are different methods. The **collocation method** determines the α_l so as to nullify the values of $L[y(x)]$ at some prescribed points x_i ; the **method of least squares** minimizes the integral of $|L[y(x)]|^2$ over the considered region; the **Galerkin method** makes $L[y(x)]$ orthogonal to every $y_l(x)$ ($l = 1, \dots, q$); and the **Ritz method** considers the variational problem $\delta J[y(x)] = 0$ whose Euler equation is $L[y(x)] = 0$ (\rightarrow 46 Calculus of Variations F), if such a problem exists, and determines the α_l so as to have $J[y(x)]$ take an extremum value with $y(x) = y_0(x) + \sum_{l=1}^q \alpha_l y_l(x)$.

References

- [1] J. D. Lambert, Computational methods in ordinary differential equations, Wiley, 1973.
- [2] C. W. Gear, Numerical initial value problems in ordinary differential equations, Prentice-Hall, 1971.
- [3] P. Henrici, Discrete variable methods in ordinary differential equations, Wiley, 1962.
- [4] G. Hall and J. M. Watt (eds.), Modern numerical methods for ordinary differential equations, Clarendon Press, 1976.
- [5] H. J. Stetter, Analysis of discretization methods for ordinary differential equations, Springer, 1970.
- [6] L. Lapidus and J. H. Seinfeld, Numerical solution of ordinary differential equations, Academic Press, 1971.
- [7] A. Ralston and P. Rabinowitz, A first course in numerical analysis, McGraw-Hill, second edition, 1978.
- [8] M. Tanaka, T. Arakawa, and S. Yamashita, On the characteristics of Runge-Kutta methods, Information Processing in Japan, 17 (1977), 175–181.
- [9] E. B. Shanks, Solutions of differential equations by evaluations of functions, Math. Comp., 20 (1966), 21–38.
- [10] K. R. Jackson, W. H. Enright, and T. E. Hull, A theoretical criterion for comparing Runge-Kutta formulas, SIAM J. Numer. Anal., 15 (1978), 618–641.
- [11] M. Tanaka, On Kutta-Merson process and its allied processes, Information Processing in Japan, 8 (1968), 44–52.
- [12] J. H. Verner, Explicit Runge-Kutta methods with estimates of the local truncation error, SIAM J. Numer. Anal., 15 (1978), 772–790.
- [13] H. Shintani, Two step processes by one step methods of order 3 and order 4, J. Sci. Hiroshima Univ., ser. A-I, 30 (1966), 183–195.
- [14] P. Merluzzi and C. Brosilow, Runge-Kutta integration algorithms with built-in estimates of the accumulated truncation error, Computing, 20 (1978), 1–16.
- [15] G. D. Byrne and R. J. Lambert, Pseudo-Runge-Kutta methods involving two points, J. Assoc. Comput. Math. 13 (1966), 114–123.
- [16] H. J. Stetter, Global error estimation in Adams PC codes, ACM Trans. Math. Software, 5 (1979), 415–430.
- [17] L. F. Shampine and M. K. Gordon, Computer solution of ordinary differential equations, the initial value problem, Freeman, 1975.
- [18] F. T. Krogh, A variable-step variable-order multistep method for the numerical solution of ordinary differential equations, Proc. Information Processing 68, North-Holland, 1969, vol. 1.

- [19] W. B. Gragg, On extrapolation algorithms for ordinary initial value problems, *SIAM J. Numer. Anal.*, 2 (1965), 384–403.
- [20] R. Bulirsch and J. Stoer, Numerical treatment of ordinary differential equations by extrapolation methods, *Numer. Math.*, 8 (1966).
- [21] G. Dahlquist, Stability questions for some numerical methods for ordinary differential equations, *Amer. Math. Soc. Proc. Symp. in Applied Math.* 15, 1963.
- [22] M. Iri, A stabilizing device for unstable numerical solutions of ordinary differential equations—Design principle and applications of a “filter,” *Information Processing in Japan*, 4 (1964), 65–73.
- [23] T. Y. Li and J. A. Yorke, Period three implies chaos, *Amer. Math. Monthly*, 82 (1975), 985–992.
- [24] M. Yamaguti and H. Matano, Euler’s finite difference scheme and chaos, *Proc. Japan Acad.*, ser. A, 55 (1979), 78–80.
- [25] A. K. Aziz (ed.), *Numerical solutions of boundary-value problem*, Academic Press, 1975.
- [26] H. B. Keller, *Numerical methods for two-point boundary-value problems*, Blaisdell, 1968.
- [27] H. B. Keller, *Numerical solutions of two-point boundary-value problems*, *SIAM Regional Conf. Ser. in Appl. Math.*, 1976.

304 (XV.9) Numerical Solution of Partial Differential Equations

A. General Remarks

The numerical solution of partial differential equations became practical in the 1950s with the advent of automatic digital computers. Nowadays, by means of modern high-performance computers, the numerical solution of partial differential equations is carried out extensively and often on a very large scale for problems in physics, engineering, and other fields of applied analysis, in order to obtain approximate solutions of rigorous equations or to simulate real phenomena by means of numerical experiments. Various numerical-solution schemes have been proposed, applied, and studied, and programming techniques required to implement numerical solutions are being invented regularly; some are of a general nature, and others are devised for particular problems. The variational method, which includes the Ritz-Galerkin method and the finite element method, and

the difference method deserve special mention here because of their generality and scope of application.

B. The Ritz-Galerkin Method

Having been applied mainly to boundary value problems for elliptic partial differential equations, the **Ritz method** is a classical method, which has been used since the old era of mechanical calculators, and is mathematically nothing other than the †direct method in the calculus of variations (→ 46 Calculus of Variations F) applied to †variational problems that are equivalent to the original boundary value problems. In order to exemplify the method, let Ω be a bounded domain in \mathbb{R}^N with piecewise smooth boundary S , and consider the Dirichlet boundary value problem comprised of the Poisson equation

$$\Delta u = -f \quad (1)$$

and the boundary condition

$$u|_S = \beta. \quad (2)$$

Here f and β are given functions on Ω and S , respectively. (In the following, given functions are assumed to be sufficiently smooth.) The boundary value problem (1), (2) is equivalent to the variational problem of minimizing the †functional

$$J[u] = \frac{1}{2}a(u, u) - (u, f) \quad (3)$$

within the set $D(J)$ of **admissible functions** u subject to (2). In (3) the following notation is used:

$$a(u, v) = (\nabla u, \nabla v)_{L_2(\Omega)} = \int_{\Omega} \nabla u \cdot \nabla v \, dx$$

and

$$(u, v) = (u, v)_{L_2(\Omega)} = \int_{\Omega} uv \, dx,$$

while

$$D(J) = \{u \in L_2(\Omega) \mid \nabla u \in L_2(\Omega), \quad u|_S = \beta\}.$$

This variational problem is further reduced to the condition:

$$u \in D(J), \quad a(u, \varphi) = (f, \varphi) \quad (\forall \varphi \in V), \quad (4)$$

where V stands for the †Sobolev space $H_0^1(\Omega)$ (→ 168 Function Spaces B), that is,

$$V = H_0^1(\Omega) = \{u \in L_2(\Omega) \mid \nabla u \in L_2(\Omega), \quad u|_S = 0\}.$$

Note that V is equal to $D(J)$ with $\beta \equiv 0$. The condition (4) is often called a **weak form of the boundary value problem**. In applying the Ritz

method to this problem, we introduce unknown parameters $\alpha_1, \alpha_2, \dots, \alpha_n$ and set the approximate solution u_n in the form

$$u_n = F(x; \alpha_1, \alpha_2, \dots, \alpha_n) \in D(J)$$

and minimize $J[u_n]$ as a function of the n variables $\alpha_1, \alpha_2, \dots, \alpha_n$. A standard way to form u_n is to choose first a $b \in D(J)$ subject to the inhomogeneous boundary condition and $\varphi_1, \varphi_2, \dots, \varphi_n \in V$ subject to homogeneous boundary condition, and then to set

$$u_n = b + \alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \dots + \alpha_n \varphi_n. \quad (5)$$

The functions φ_j ($j = 1, 2, \dots, n$) are called **basis functions** or **coordinate functions** in the Ritz method. Denoting by D_n the set of all possible functions appearing on the right-hand side of (5) and by V_n the linear space generated by $\varphi_1, \varphi_2, \dots, \varphi_n$, we can state the condition to determine the approximate solution u_n as:

$$u_n \in D_n$$

$$a(u_n, \varphi) = (f, \varphi) \quad (\forall \varphi \in V_n). \quad (6)$$

Thus the Ritz method is a kind of **projective approximation method** which approximates the weak equation (4) by its projection on a finite-dimensional space V_n . The condition (6) is equivalent to the equations

$$a(u_n, \varphi_j) = (f, \varphi_j) \quad (j = 1, 2, \dots, n). \quad (7)$$

Particularly when $\beta \equiv 0$ in (2), D_n coincides with V_n , and the equations (7) that determine the coefficients α_j are reduced to a linear equation

$$K\alpha = \gamma \quad (8)$$

for $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, where the matrix $K = (K_{ij})$ and the vector $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$ are given respectively by $K_{ij} = a(\varphi_i, \varphi_j)$ and $\gamma_j = (f, \varphi_j)$. Solving (8) numerically by the †Gauss elimination method or by †iteration, one eventually obtains the approximate solution u_n by the Ritz method. The Ritz method is also applicable, for instance, to the eigenvalue problem $\Delta u = \lambda u, u|_s = 0$, since this eigenvalue problem is equivalent to the variational problem of finding a stationary **Rayleigh's quotient** $R[u] = a(u, u)/(u, u)$ with V as the set of admissible functions. In the approximation to restrict the admissible functions to V_n , the approximate eigenvalue $\lambda^{(n)}$ is determined through the matrix eigenvalue problem

$$K\alpha = \lambda^{(n)} M\alpha, \quad (9)$$

where the matrix K is as above and the matrix $M = (M_{ij})$ is given by $M_{ij} = (\varphi_i, \varphi_j)$. Sometimes, optimization of $J[u_n]$ or $R[u_n]$ is carried out more directly, e.g., by the †gradient method, particularly when u_n contains free parameters α_j in some nonlinear way.

Many studies have been made of convergence and error estimation for the Ritz method in general (\rightarrow e.g., [1–4]). For instance, it is known that, if $\beta = 0$, $f \in L^2(\Omega)$, and $\{\varphi_1, \varphi_2, \dots, \varphi_n, \dots\}$ spans a linear subspace dense in $V = H_0^1(\Omega)$, then the approximate solution u_n obtained through (6) (or through (12) below) converges to the rigorous solution in the H^1 -topology (and hence in the L^2 -topology). Nevertheless, success in using the method in practical applications can be gained only by a clever choice of trial functions. A typical but systematic example of such a choice is made by the finite element method (\rightarrow Section C).

We proceed to the **Galerkin method** [1]. Suppose that we are to solve the equation

$$L[u] = \Delta u + (b \cdot \nabla)u = -f, \quad (10)$$

which is the equation in (1) perturbed by a lower-order term (the convection term). Here $b = (b_1(x), \dots, b_n(x))$ and $(b \cdot \nabla)u = \sum_j b_j(x)(\partial u / \partial x_j)$. For simplicity, the boundary condition (2) is assumed to be homogeneous (i.e., $\beta \equiv 0$). Although the boundary value problem (10), (2) cannot be reduced to a variational problem since (10) is not symmetric, it is equivalent to its weaker form

$$a(u, \varphi) - ((b \cdot \nabla)u, \varphi) = (f, \varphi) \quad (\forall \varphi \in V), \quad (11)$$

provided that $u \in V$. Note that (11) is obtained from $(L[u] + f, \varphi)_{L_2} = 0$ by transforming $(L[u], \varphi)$ through integration by parts. Actually, this boundary value problem has a unique solution u for any $f \in L_2(\Omega)$ if $|\operatorname{div} b|$ is sufficiently small. According to the Galerkin method as applied to the present problem, one replaces V by V_n in (11) and determines the approximate solution $u_n = \alpha_1 \varphi_1 + \dots + \alpha_n \varphi_n \in V_n$ by the condition

$$a(u_n, \varphi) - ((b \cdot \nabla)u_n, \varphi) = (f, \varphi) \quad (\forall \varphi \in V_n), \quad (12)$$

which is equivalent to the equations

$$a(u_n, \varphi_j) - ((b \cdot \nabla)u_n, \varphi_j) = (f, \varphi_j) \quad (j = 1, 2, \dots, n). \quad (13)$$

Sometimes it is convenient to project $(L[u] + f, \varphi) = 0$ onto the finite-dimensional space and to deal directly with $(L[u_n] + f, \varphi_j) = 0$ ($j = 1, 2, \dots, n$). When the coefficients of the equation and the given function f are periodic in space variables and when sine or cosine functions are chosen as basis functions, the Galerkin method turns out to be the same as the so-called Fourier approximation method and can be efficiently implemented with the aid of a †fast Fourier transform (FFT) to yield the required Fourier coefficients numerically [5].

The Galerkin method can be applied to the numerical solution of evolution equations, namely, in order to approximate time-

dependent solutions. This is exemplified through the following initial-boundary value problem for the diffusion equation (heat equation):

$$\frac{\partial u}{\partial t} = \Delta u \quad (t \geq 0, x \in \Omega) \quad (14)$$

with the homogeneous boundary condition

$$u|_s = 0 \quad (15)$$

and the initial condition

$$u|_{t=0} = a(x) \quad (x \in \Omega). \quad (16)$$

We first set the required approximate solution $u_n = u_n(t, x)$ in terms of the basis functions $\varphi_1, \varphi_2, \dots, \varphi_n$ of V_n as

$$u_n = \alpha_1(t)\varphi_1 + \alpha_2(t)\varphi_2 + \dots + \alpha_n(t)\varphi_n, \quad (17)$$

and then determine the coefficients $\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)$ from the conditions

$$\frac{d}{dt}(u_n, \varphi) = -a(u_n, \varphi) \quad (t \geq 0, \forall \varphi \in V_n) \quad (18)$$

and

$$(u_n(0), \varphi) = (a, \varphi) \quad (\forall \varphi \in V_n). \quad (19)$$

Since (18) is equivalent to the equations

$$\frac{d}{dt}(u_n, \varphi_j) = -a(u_n, \varphi_j) \quad (j = 1, 2, \dots, n),$$

the vector function $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t))$ is obtained by solving the ordinary differential equation

$$M \frac{d}{dt} \alpha = -K \alpha, \quad (20)$$

where the matrices M and K are those in (8) and (9). The initial value $\alpha(0)$ of α should be given in accordance with (19). Sometimes the procedure above is called a **semidiscrete approximation** based on the Galerkin method. Obviously, this method can be applied to more general evolution equations, for instance, to the **diffusion-convection equation**

$$\frac{\partial u}{\partial t} = L[u] = \Delta u + (b \cdot \nabla)u, \quad (21)$$

which governs time-dependent states corresponding to stationary states subject to (10). As a concrete example, we can refer to the finite element approximation for evolution equations described in Section D. If the geometry of Ω is simple enough, e.g., for 1-dimensional Ω , one can adopt eigenfunctions of $L[\]$ as the basis functions; the resulting version of the semidiscrete Galerkin method is called a **spectral method**.

In order to obtain $u_n(t)$ one has to carry out numerical integration for (20), discretizing the time variable. The Galerkin approxima-

tion with discretization of the space and time variables is called a **full discrete approximation**. Finally, an advantage of the Ritz-Galerkin method is that if the boundary condition imposed on rigorous solutions is a natural boundary condition, then the admissible functions and particularly the basis functions need not satisfy it [4].

C. Finite Element Methods for Boundary Value Problems

In recent years the method of numerical solution of partial differential equations most extensively employed in structural mechanics and many other fields of engineering has been the **finite element method**. In its standard form, the finite element method can be regarded, at least mathematically, as a type of Ritz-Galerkin method that adopts as its basis functions piecewise polynomials of low degree and with narrow supports. Although the idea leading to the finite element method can be traced back to a paper by R. Courant in 1943 (*Bull. Amer. Math. Soc.*), the method acquired its popularity in the late 1950s when it was rediscovered by engineers on the basis of mechanical considerations [6]. Here we apply the method to the boundary value problem (1), (2), assuming that $\beta = 0$ and Ω a polygonal domain in \mathbf{R}^2 [7–9]. First, $\Omega = \Omega \cup S$ is divided into small triangles of which the length of any side does not exceed $h > 0$. Each triangle T appearing in this decomposition of $\bar{\Omega}$ is called a **triangular element**, or simply an **element** (following the terminology in structural mechanics), and we denote the set of all triangular elements by \mathcal{T}_h , h being equal to $\max_{T \in \mathcal{T}_h} \{ \text{the longest side of } T \}$. The set of all **nodal points**, i.e., the vertices of the triangular elements, is denoted by N , and we put

$$N_i = \{ P \in N \mid P \in \Omega \} = \{ P_1, P_2, \dots, P_n \}. \quad (22)$$

Then a standard choice of admissible functions is to adopt as V_n in (6) the following $V_h \subset V = H_0^1(\Omega)$:

$$V_h = \{ v_h \in C(\bar{\Omega}) \mid v_h \text{ is linear on each element} \\ \text{and } v_h|_s = 0 \}, \quad (23)$$

namely, V_h is the set of all elementwise linear continuous functions satisfying the boundary condition. Now the approximate solution $u_h \in V_h$ is determined by the condition

$$a(u_h, \varphi_h) = (f, \varphi_h) \quad (\forall \varphi_h \in V_h). \quad (24)$$

A basis of V_h is formed by pyramidal functions φ_j ($j = 1, 2, \dots, n$) $\in V_h$ such that

$$\varphi_j(P_j) = 1 \quad \text{and} \quad \varphi_j(Q) = 0 \quad (Q \in N \setminus \{P_j\}). \quad (25)$$

In order to obtain u_h , one has to solve (8) numerically for α . Since the support of φ_j is confined to elements adjacent to P_j , the matrix K turns out to be a sparse matrix of **multi-diagonal type**, and hence particular elimination procedures can be applied with high efficiency. In the finite element method, the matrix K is often called the **stiffness matrix**. The convergence of u_h to u as $h \rightarrow 0$ is guaranteed if \mathcal{T}_h ($0 < h \leq h_0$) satisfies a certain regularity condition. For instance, if any angle of a triangle T is not smaller than $\theta_0 > 0$, then u_h converges in the $H^1(\Omega)$ -topology. Moreover, when Ω is convex, it holds that

$$\begin{aligned} \|u - u_h\|_{H^1} &\leq Ch \|f\|_{L_2}, \\ \|u - u_h\|_{L_2} &\leq Ch^2 \|f\|_{L_2} \end{aligned} \quad (26)$$

[7, 9]. Uniform convergence of u_h to u and L_∞ -estimation of the error have also been established (e.g., [7, chap. 3]). u_h is subject to the maximum principle, provided that the triangulation \mathcal{T}_h is of acute type. Here, \mathcal{T}_h is said to be of **acute type** (or of **strongly acute type**) if any angle θ of $T \in \mathcal{T}_h$ is in $0 < \theta \leq (\pi/2)$ ($0 < \theta \leq \theta_1 < (\pi/2)$). To acquire higher accuracy, more sophisticated admissible functions that are piecewise polynomials of higher degree are used. To take account of curved boundaries, several devices, including the **isoparametric method**, have been proposed.

In dealing with biharmonic equations or other equations of higher order, one can conveniently apply a modified version of the finite element method of **mixed type** or of **nonconforming type**; in the latter the approximate solution is sought within a class of functions which are of less regularity at the interface of elements and hence do not belong to the domain of the original variational problem. i.e., which are not admissible in the classical sense [10, 11].

D. Finite Element Methods for Initial Value Problems

The most basic type of finite element method applicable to evolution problems is described here, and concerns the initial boundary value problem (14)–(16), assuming that Ω is the same as in Section B [8, 9]. In the semidiscrete finite element approximation one uses the functions φ_j of the preceding section as the basis for the approximate solution, now denoted by u_h , and then determines u_h through (17)–(19) with V_n replaced by V_h . Then u_h converges to u as $h \rightarrow 0$, provided that the triangulation \mathcal{T}_h is regular. Furthermore, under the same condition that makes (26) hold, the error is given by the fol-

lowing estimate [12]:

$$\|u - u_h\|_{L_2} \leq Ch^2 \|a\|_{L_2/t} \quad (t > 0), \quad (27)$$

which reflects the smoothing property of the diffusion equation. In the terminology of the finite element method, the matrix M in (20) is called the **mass matrix**. In order to gain a fully discrete finite element approximation $u_{\tau,h}$ for u , one discretizes the time variable with the \dagger mesh length τ and solves the difference analog of (20), i.e., either

$$M(\alpha(t + \tau) - \alpha(t))/\tau = -K\alpha(t) \quad (28)$$

or

$$M(\alpha(t + \tau) - \alpha(t))/\tau = -K\alpha(t + \tau), \quad (29)$$

where t is restricted to $t = k\tau$ ($k = 0, 1, 2, \dots$).

The difference schemes (28) and (29) are of **forward type** and of **backward type**, respectively. From (29) follows $\|\alpha(k\tau)\| \leq \|\alpha(0)\|$, and consequently $\|u_{\tau,h}(t)\|_{L_2} \leq C\|a\|_{L_2}$. In the generally accepted terminology, the approximation is **stable** if for each $T > 0$ there exists a positive constant C_T depending on T such that

$$\|u_{\tau,h}(t)\| \leq C_T \|a\| \quad (0 \leq t = k\tau \leq T). \quad (30)$$

Thus the approximation through (29) is unconditionally stable, while the one through (28) is stable only under certain conditions on the triangulation \mathcal{T}_h and on the time mesh τ . For instance, the forward scheme (28) is stable if \mathcal{T}_h is regular and satisfies the **inverse assumption**

$$\sup_{0 < h \leq h_0} \max_{T \in \mathcal{T}_h} \frac{h}{(\text{longest side of } T)} < +\infty,$$

and if the ratio τ/h^2 is sufficiently small [13, 14].

Sometimes, the mass matrix M in (28) and (29) is modified by the following procedure, called **mass lumping**: For each P_j , one joins alternately the center of gravity of the elements with P_j as one of their vertices and the middle point of the sides with P_j as one of their endpoints, thus forming a closed broken line Γ_j surrounding P_j . We denote by $\bar{\varphi}_j$ the characteristic function of the polygonal domain B_j bounded by Γ_j . When mass lumping is applied, the matrix M is replaced by the matrix $\bar{M} = ((\bar{\varphi}_i, \bar{\varphi}_j))$. Generally, mass lumping relaxes the conditions for the stability and the maximum principle to hold, although it may lower the accuracy to some extent (H. Fujii, *Proc. US-Japan Seminar*, Univ. Tokyo Press, 1973). Usually, schemes without mass lumping are called **consistent mass schemes**.

In principle, there is no difficulty in applying the finite element method to the diffusion-convection equation (21). However, when $\|b\|$ is large, the stability condition becomes strin-

gent. To meet these difficulties, M. Tabata, T. Ikeda [15], and others have devised some schemes that enjoy better stability, the maximum principle, and even the law of conservation of mass by improving mass lumping and introducing the idea of an upstream approximation or an artificial viscosity for discretizing the convection term. Moreover, finite element methods are currently being applied to wave equations, the Navier-Stokes equations, and various other evolution equations.

E. Difference Methods for Boundary Value Problems

In difference methods for solving partial differential equations we reduce the original equation to its **difference analog**, replacing differential quotients of the unknown function by the corresponding difference quotients. The difference analog, usually a system of algebraic equations, is then solved numerically, yielding the desired approximate solution. Since the last part of the method, i.e., the numerical solution of the difference analog, requires a large amount of computation, difference methods were not feasible until the development of automatic computers.

In terms of the variable x , difference approximations for the derivative df/dx are the **forward difference**

$$(D_x^h f)(x) = (f(x+h) - f(x))/h,$$

the **backward difference**

$$(D_x^h f)(x) = (f(x) - f(x-h))/h,$$

and the **central difference (symmetric difference)**

$$(\delta_x^h f)(x) = (f(x+h) - f(x-h))/2h.$$

In order to exemplify the difference method applied to boundary value problems, we return to equations (1) and (2), supposing that Ω is a 2-dimensional square $Q = \{(x, y) | 0 < x < 1, 0 < y < 1\}$ [1, 16–18]. Then we cover $Q \cup S$ by a **square net** (lattice, grid) with mesh length $h = 1/N$, for N a large integer. The set of net points $P_{m,n} = (mh, nh)$ lying in Q is denoted by Ω_h , and S_h is the set of net points on S . We want to have a net function u_h on $R_h = \Omega_h \cup S_h$ that approximates the solution u of (1) and (2), where u_h is determined by the difference equation

$$\Delta_h u_h(x, y) = -f(x, y) \quad ((x, y) \in \Omega_h), \quad (31)$$

and the boundary condition

$$u_h(x, y) = \beta(x, y) \quad ((x, y) \in S_h). \quad (32)$$

Here Δ_h is a difference operator, the 5-point

difference analog of Δ , which is defined by

$$\begin{aligned} \Delta_h \varphi &= D_x^h D_x^h \varphi + D_y^h D_y^h \varphi \\ &= (\varphi(x+h, y) + \varphi(x-h, y) + \varphi(x, y+h) \\ &\quad + \varphi(x, y-h) - 4\varphi(x, y))/h^2. \end{aligned}$$

In view of the maximum principle for net functions on R_h , it is seen that (31) and (32) admit only $u_h = 0$ as a solution if $f = \beta = 0$, which implies that the solution u_h of (31) and (32) exists uniquely for all f and β . When we are concerned with the convergence of u_h to u as $h \rightarrow 0$, we have to consider (31) and (32) for all small $h > 0$. Such a family of difference equations is called a **difference scheme**. Actually, u_h obtained through (31) and (32) converges to u as $h \rightarrow 0$, and we can say that the difference scheme (31) and (32) is convergent. In general, the order of the error $u - u_h$ depends on the smoothness of u . For instance, if $u \in C^4(\bar{\Omega})$, then $|u - u_h| \leq Ch^2$ holds uniformly on R_h . When the asymptotic behavior of the error is known to satisfy $u - u_h = wh^2 + o(h^2)$ for some $w = w(x, y)$, it is possible to attain a higher accuracy by eliminating the h^2 -order term if we compute u_h for two different values of h and form an appropriate linear combination of the u_h thus obtained (Richardson's extrapolation; \rightarrow e.g., [17]). The convergence of difference schemes applied to boundary value problems is not affected if rectangular nets are used instead of square nets. Many techniques have been devised to treat curved boundaries. There have been attempts recently to map Ω into a domain Q of simpler geometry, say, a rectangular one, and then discretize the arising partial differential equations with variable coefficients in Q (e.g., J. F. Thompson et al., *J. Comput. Phys.*, 1982.) Moreover, when information regarding the location and nature of singularities of solutions is available, one can refine the net near these points.

The difference analog of elliptic equations described above is a system of linear equations with the values of the u_h as its unknowns. Since the coefficient matrix of this system is of multidagonal type, direct elimination methods can be efficiently applied to solve it. Particularly, some elimination algorithms employing vector computations have been invented [19]. Iteration methods of the Gauss-Seidel type can also be applied, where acceleration of convergence by SOR is effective (\rightarrow 302 Numerical Solution of Linear Equations C).

F. Difference Methods for Initial Value Problems

To demonstrate characteristics of the difference method applied to evolution equa-

tions [17, 20, 21], we first consider the initial boundary value problem (14)–(16) for the diffusion equation supposing that $\Omega = (0, 1)$, a 1-dimensional interval. Then we cover $Q = [0, \infty) \times \bar{\Omega}$ by a rectangular net $t = n\tau$ ($n = 0, 1, 2, \dots$) and $x = jh$ ($j = 0, 1, \dots$) with mesh lengths $\Delta t = \tau$ and $\Delta x = h$. The value of the approximate function $u_{\tau,h}$ at the net point $(n\tau, jh)$ is denoted by U_j^n . Then the simplest **difference scheme of forward type** for (14) is written as

$$\frac{U_j^{n+1} - U_j^n}{\tau} = \frac{U_{j+1}^n + U_{j-1}^n - 2U_j^n}{h^2}. \quad (33)$$

Namely, we adopt the difference operator $L_{\tau,h}[\] = D_t - D_x^h D_x^h$ as the difference analog of $(\partial/\partial t) - (\partial^2/\partial x^2)$. Equation (33) is **consistent** in the sense that $L_{\tau,h}[u] = \text{"the residual of } u\text{"} \rightarrow 0$ ($\tau, h \rightarrow 0$) for any smooth solution u of (14). This is the case also with the **difference scheme of backward type**

$$\frac{U_j^{n+1} - U_j^n}{\tau} = \frac{U_{j+1}^{n+1} + U_{j-1}^{n+1} - 2U_j^{n+1}}{h^2}. \quad (34)$$

To compute the approximate solution $U^n = \{U_j^n | 0 \leq j \leq N\}$, (33) or (34) must be combined with the approximate initial condition corresponding to (16), say $U_j^0 = a(jh)$, and with the boundary condition $U_0^n = U_N^n = 0$. Then one can proceed from U^0 to U^1 , from U^1 to U^2 , and so on. Actually, (33) is rewritten in the form

$$U_j^{n+1} = \lambda U_{j+1}^n + \lambda U_{j-1}^n + (1 - 2\lambda)U_j^n, \quad (35)$$

with $\lambda = \tau/h^2$, which gives U^{n+1} explicitly in terms of U^n . Hence (33) is an **explicit scheme**. On the other hand, in order to acquire U^{n+1} from U^n according to (34), it is necessary to solve a system of linear equation with U^{n+1} as its unknown. Hence (34) is an **implicit scheme**. Since the coefficient matrix of the system of linear equations to determine U^{n+1} according to (34) is tridiagonal, one can employ a particularly efficient elimination algorithm.

We introduce the maximum norm $\|\varphi_h\|_h = \max_{R_h} |\varphi(jh)|$ for net functions φ_h on $R_h = \Omega_h \cup S_h = \{x = jh | 0 \leq j \leq N\}$. Then, as is obvious from (35), the approximate solution $U^n = u_{\tau,h}(n\tau)$ obtained through (33) satisfies $\|U^n\|_h \leq \|U^0\|_h$ ($n = 0, 1, \dots$), provided that

$$0 < \lambda = \frac{\tau}{h^2} \leq \frac{1}{2}. \quad (36)$$

Generally, a difference scheme approximating an initial value problem is said to be **stable** if the approximate solution $u_{\tau,h}$ for the initial value $a_h = u_{\tau,h}(0)$ satisfies, for small τ and h ,

$$\|u_{\tau,h}(t)\|_h \leq M_T \|a_h\|_h \quad (0 \leq t = n\tau \leq T) \quad (37)$$

for any $T > 0$, M_T being a constant depending

on T . **Lax's equivalence theorem** asserts that if the original initial value problem is well-posed and if the approximating difference scheme is consistent, then the stability of the difference scheme is necessary and sufficient for the convergence of the approximate solution $u_{\tau,h}$ to the rigorous solution u as $\tau, h \rightarrow 0$. For instance the explicit difference scheme (33) for the 1-dimensional diffusion equation is convergent under the mesh condition (36). On the other hand, the implicit scheme (34) is unconditionally stable and hence is convergent. Many difference schemes more sophisticated than (33) and (34), of higher accuracy or with other favorable properties, have been proposed and studied. All these difference schemes can be generalized to the case of many-dimensional Ω . For instance, if $\Omega \subset \mathbf{R}^2$, then by means of the difference operator Δ_h in (31), the forward scheme and the backward scheme are given by $D_t U^n = \Delta_h U^n$ and $D_t U^n = \Delta_h U^{n+1}$, respectively. The former is stable if $0 < \lambda = \tau/h^2 \leq 1/4$, while the latter is unconditionally stable. Moreover, in a method called the **ADI method** (alternating direction implicit method) one introduces $U^{n+1/2}$ on **fractional steps** $t = n\tau + \tau/2$ and discretizes (14) as

$$\begin{aligned} \frac{U^{n+1/2} - U^n}{\tau/2} &= D_x^h D_x^h U^n + D_y^h D_y^h U^{n+1/2}, \\ \frac{U^{n+1} - U^{n+1/2}}{\tau/2} &= D_x^h D_x^h U^{n+1/2} + D_y^h D_y^h U^{n+1/2}. \end{aligned} \quad (38)$$

Then, starting from U^0 , the computation goes as $U^0 \rightarrow U^{1/2} \rightarrow U^1 \rightarrow \dots \rightarrow U^n \rightarrow U^{n+1/2} \dots$ by solving at each step alternatively a difference equation implicit with respect to x and y , for which a particular elimination method can be used with good efficiency. Furthermore, the ADI method is unconditionally stable. Various types of fractional-step method that generalize the ADI method have been proposed [12, 17].

We now proceed to hyperbolic equations. First consider the Cauchy problem where the spatial domain is the whole \mathbf{R}^1 . Noting that the wave equation $v_{tt} = v_{xx}$, for instance, can be rewritten as

$$\frac{\partial u}{\partial t} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial u}{\partial x}$$

by putting $u = {}^t(v_t, v_x)$, we here deal with a hyperbolic system of the form

$$\frac{\partial u}{\partial t} = A \frac{\partial u}{\partial x}, \quad (39)$$

where the unknown u is an m -vector $u = {}^t(u_1, u_2, \dots, u_m)$, and A is a constant $m \times m$ matrix that is supposed to be real and symmetric. For hyperbolic equations, explicit schemes are generally preferred, and as a typ-

ical one we mention the **Friedrichs scheme**. In the Friedrichs scheme applied to (39), the approximate solution $U = {}^t(U_1, U_2, \dots, U_m)$ is determined by

$$\frac{U(t + \tau, x) - \frac{1}{2}(U(t, x + h) + U(t, x - h))}{\tau} = A \delta_x^h U(t, x). \quad (40)$$

Here t is restricted to $t = n\tau$ ($n = 0, 1, \dots$), while x is regarded as a variable ranging over \mathbf{R}^1 . Introducing translation operators T_h and T_{-h} by

$$T_h: \varphi(x) \rightarrow \varphi(x + h) \quad \text{and} \quad T_{-h}: \varphi(x) \rightarrow \varphi(x - h),$$

we obtain from (40)

$$U(t + \tau, \cdot) = S_h U(t, \cdot), \quad (41)$$

where, with the mesh ratio $\lambda = \tau/h$, S_h is given by

$$S_h = \left(\frac{1}{2} + \frac{\lambda}{2} A \right) T_h + \left(\frac{1}{2} - \frac{\lambda}{2} A \right) T_{-h}. \quad (42)$$

Generally, if a difference scheme is reduced to the form of (41), then the operator S_h , which yields evolution of the solution of the difference equation for one time step, is called the **amplification operator** of the scheme. The stability of the difference scheme is implied by the uniform boundedness of the operator norm $\|S_h^n\|$ of S^n acting in the Hilbert space $(L_2(\mathbf{R}^1))^m$ for $0 \leq n \leq T/\tau$. Consequently, the scheme is stable if $\|S_h\| \leq 1 + C\tau$, C being a constant. The symbol $\tilde{S}_h(\xi)$ of S_h with respect to the Fourier transform is called the **amplification matrix** of the scheme. The L_2 -stability of the scheme mentioned above is equivalent to the uniform boundedness of the matrix norm $\|\tilde{S}_h^n(\xi)\|$ for $0 \leq n\tau \leq T$ and $\xi \in \mathbf{R}^1$. Therefore, denoting by $r_h(\xi)$ the spectral radius of $\tilde{S}_h(\xi)$, one can give a necessary condition for the stability by $|r_h(\xi)| \leq 1 + C\tau$, which is called the **von Neumann condition**. Concerning general criteria for the uniform boundedness of powers of $\tilde{S}_h(\xi)$ when $\tilde{S}_h(\xi)$ is not necessarily normal, a fundamental theorem was given by H. O. Kreiss in 1962 [21].

If the largest-modulus eigenvalue of A is denoted by μ_0 , then the Friedrichs scheme is stable and hence is convergent under the condition on the mesh ratio $\lambda = \tau/h$ given by

$$\lambda \leq 1/|\mu_0|. \quad (43)$$

In particular, (43) implies that for the stable Friedrichs scheme

$$\text{propagation speed of the difference scheme} \geq \text{propagation speed of the original equation.} \quad (44)$$

Condition (44) is a necessary condition for general difference schemes approximating

hyperbolic equations to be stable, and is known as the **CFL condition** (Courant-Friedrichs-Lewy condition) [22]. For a scheme more accurate than the Friedrichs scheme we refer to the Lax-Wendroff scheme, which proceeds by way of the following amplification operator S_h :

$$\begin{aligned} S_h &= I + \tau A \delta_x^h + \frac{\tau^2}{2} A^2 D_x^h D_x^h \\ &= I + \frac{\lambda}{2} A (T_h + T_{-h}) + \frac{\lambda^2}{2} (T_h - 2I + T_{-h}). \end{aligned}$$

The Lax-Wendroff scheme is stable under (43).

There are many other difference schemes to approximate (39), most of which are applicable to higher-dimensional spaces. Some of these schemes can be conveniently employed to solve nonlinear hyperbolic equations, for instance, those arising in the gas dynamics of compressible fluids, for which the Friedrichs scheme and the Lax-Wendroff scheme were originally intended [23]. Concerning difference schemes with variable coefficients which approximate a regularly hyperbolic system with an x -dependent principal part, criteria for the stability of the scheme in terms of the symbol $\tilde{S}_h(x, \xi)$ of S_h have been obtained, making use of the theory of pseudodifferential operators, by P. D. Lax and L. Nirenberg (*Comm. Pure Appl. Math.*, 1966), M. Yamaguti and T. Nogi (*Publ. Res. Inst. Math. Sci.*, 1967), R. Vaillancourt (*Math. Comput.*, 1970), Z. Koshiba (*J. Math. Anal. Appl.*, 1981), and others.

References

- [1] L. V. Kantorovich and V. I. Krylov, Approximate methods of higher analysis, Interscience and Noordhoff, 1958. (Original in Russian, 1950.)
- [2] J. Necas, Les méthodes directes en théorie des équations elliptiques, Masson, 1967.
- [3] R. Temam, Numerical analysis, Reidel, 1973.
- [4] R. Courant and D. Hilbert, Methods of mathematical physics, Interscience, I, 1953; II, 1962.
- [5] J. W. Cooley and J. W. Tukey, An algorithm for the machine calculation of complex Fourier Series, *Math. Comp.*, 19 (1964), 297–301.
- [6] M. J. Turner, R. W. Clough, H. C. Martin, and L. J. Topp, Stiffness and deflection analysis of complex structures, *J. Aero. Sci.*, 23 (1956), 805–823.
- [7] P. G. Ciarlet, The finite element method for elliptic problems, North-Holland, 1978.
- [8] A. R. Mitchell and R. Wait, The finite element method in partial differential equations, Wiley, 1977.

- [9] G. Strang and G. Fix, An analysis of the finite element method, Prentice-Hall, 1973.
- [10] F. Kikuchi and Y. Ando, On the convergence of a mixed finite scheme for plate bending, *Nuclear Eng. Design*, 24 (1973), 357–373.
- [11] T. Miyoshi, A mixed finite element method for the solution of the von Karman equations, *Numer. Math.*, 26 (1976), 255–269.
- [12] H. Fujita and A. Mizutani, On the finite element method for parabolic equations I, Approximation of holomorphic semi-groups, *J. Math. Soc. Japan*, 28 (1976), 749–771.
- [13] T. Kato, Perturbation theory for linear operators, Springer, second edition, 1976.
- [14] T. Ushijima, Approximation theory for semi-groups of linear operators and its application to approximation of the wave equation, *Japan. J. Math.*, 1 (1975), 185–224.
- [15] T. Ikeda, Maximum principle in finite element models for convection-diffusion phenomena, Kinokuniya and North-Holland, 1983.
- [16] G. E. Forsythe and W. R. Wasow, Finite-difference methods for partial differential equations, Wiley, 1960.
- [17] G. I. Marchuk, Methods of numerical mathematics, Springer, second edition, 1982.
- [18] G. D. Smith, Numerical solution of partial differential equations: Finite difference methods, Clarendon Press, second edition, 1978.
- [19] D. L. Book (ed.), Finite-difference techniques for vectorized fluid dynamics calculations, Springer, 1981.
- [20] F. John, Lectures on advanced numerical analysis, Gordon & Breach, 1967.
- [21] R. D. Richtmyer and K. W. Morton, Difference method for initial value problems, Interscience, 1957.
- [22] R. Courant, K. Friedrichs, and H. Lewy, Über die partiellen Differenzengleichungen der mathematischen Physik, *Math. Ann.*, 100 (1928), 32–74.
- [23] R. Peyret and T. D. Taylor, Computational methods for fluid flow, Springer, 1983.
- [24] B. Carnahan, H. A. Luther, and J. O. Wilkes, Applied numerical methods, Wiley, 1969.